



0410
2623
5

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In the application of:
Yoshihisa Shinagawa et al
Serial No.: 09/781,309

Filed: February 13, 2001
For: METHOD AND APPARATUS FOR
CODING AND DECODING IMAGE
DATA USING CRITICAL POINTS

Assistant Commissioner for Patents
Washington, D.C. 20231
USA

Dear Sir:

Group Art Unit: 2623

Attorney Docket: 13599

RECEIVED

MAY 24 2001

Technology Center 2600

To complete the claim for foreign priority made under 35 U.S.C. § 119(a)-(d) or (f) in respect of the above-noted application, the applicant enclosed the certified copy of Japanese Patent Application No. 2000-69842.

Respectfully submitted,

Ralph A. Dowell
Registration No. 26, 868

Dowell & Dowell P.C.
Suite 309
1215 Jefferson Davis Highway
Arlington, Virginia
U.S.A. 22202

Dated: 05/18/2001

NWH/hr 92637-12



PATENT OFFICE

JAPANESE GOVERNMENT

This is to certify that the annexed is a true copy of the following application as filed with this Office.

RECEIVED

DATE OF APPLICATION : March 14, 2000

MAY 24 2001

Technology Center 2600

APPLICATION NUMBER : Patent Application 2000-69842

APPLICANT(S) : MONOLITH CO., LTD.

March 30, 2001

Commissioner, Patent Office

Kozo OIKAWA

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application: 2000年 3月 14日

出願番号
Application Number: 特願 2000-069842

願人
Applicant(s): 株式会社モノリス



RECEIVED

MAY 24 2001

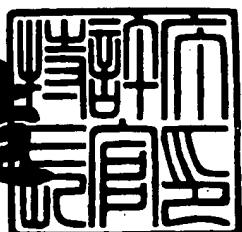
Technology Center 2600

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 3月 30日

特許庁長官
Commissioner,
Patent Office

及川耕造



【書類名】 特許願
 【整理番号】 MN-0102
 【提出日】 平成12年 3月14日
 【あて先】 特許庁長官殿
 【国際特許分類】 G06T 5/00
 G06T 7/00

【発明者】

【住所又は居所】 東京都江戸川区西葛西5-10-26
 【氏名】 品川 嘉久

【発明者】

【住所又は居所】 千葉県千葉市稲毛区天台4-2-12
 【氏名】 永嶋 浩樹

【特許出願人】

【識別番号】 396001980
 【氏名又は名称】 株式会社モノリス

【代理人】

【識別番号】 100105924
 【弁理士】
 【氏名又は名称】 森下 賢樹
 【電話番号】 0422-23-7415

【手数料の表示】

【予納台帳番号】 054977
 【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 9717729
 【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像データ符号化および復号のための方法および装置

【特許請求の範囲】

【請求項1】 画像データに含まれるフレームをキーフレームと中間フレームに分別する工程と、

分別によって得られたキーフレームの間でマッチングを計算する工程と、

マッチングの結果をもとに仮想的な中間フレームを生成する工程と、

仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する工程と、

を含むことを特徴とする画像データ符号化方法。

【請求項2】 前記計算する工程は、キーフレーム間で画素単位のマッチングを計算し、

前記生成する工程は、キーフレーム間における画素の位置および輝度の対応関係をもとに、画素単位で補間計算を実施して仮想的な中間フレームを生成する請求項1に記載の画像データ符号化方法。

【請求項3】 前記符号化する工程は、仮想的な中間フレームと現実の中間フレームの差分を符号化する請求項1、2のいずれかに記載の画像データ符号化方法。

【請求項4】 前記キーフレームのデータと前記符号化する工程で得られたデータの組合せを前記画像データの符号化データとして出力する工程をさらに含む請求項1から3のいずれかに記載の画像データ符号化方法。

【請求項5】 コンピュータにて実行可能なプログラムを格納した記録媒体であって、前記プログラムは、

画像データに含まれるフレームをキーフレームと中間フレームに分別する工程と、

分別によって得られたキーフレームの間でマッチングを計算する工程と、

マッチングの結果をもとに仮想的な中間フレームを生成する工程と、

仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する工程と、

を含むことを特徴とする、コンピュータにて読み取り可能な記録媒体。

【請求項6】 複数のフレームを含む画像データを取得するユニットと、前記画像データに含まれるフレームをキーフレームと中間フレームに分別するユニットと、

分別によって得られたキーフレームを入力し、それらキーフレーム間のマッチングを計算するユニットと、

マッチングの計算結果をもとに仮想的な中間フレームを生成するユニットと、仮想的な中間フレームを利用して前記分別によって得られた現実の中間フレームを符号化するユニットと、

を含むことを特徴とする画像データ符号化装置。

【請求項7】 前記符号化するユニットは、前記仮想的な中間フレームと前記現実の中間フレームの差分を符号化する請求項6に記載の画像データ符号化装置。

【請求項8】 前記マッチングを計算するユニットは、キーフレーム間で画素単位のマッチングを行うことを特徴とする請求項6、7のいずれかに記載の画像データ符号化装置。

【請求項9】 前記生成するユニットは、前記画素単位のマッチングの結果をもとに、前記キーフレームの画素間を補間することで、前記仮想的な中間フレームを生成する請求項6から8のいずれかに記載の画像データ符号化装置。

【請求項10】 前記生成するユニットは、前記キーフレーム間における画素の位置および輝度の対応関係をもとに、画素単位で補間計算を実施して前記仮想的な中間フレームを生成する請求項6から8のいずれかに記載の画像データ符号化装置。

【請求項11】 前記キーフレームのデータと前記符号化するユニットの出力データを組み合わせ、これを前記画像データの符号化データとして出力するユニットをさらに含む請求項6から10のいずれかに記載の画像データ符号化装置。

【請求項12】 画像データの符号化データに含まれる前記画像データのキーフレームとその他の付加データを分別する工程と、

分別によって得られたキーフレームの間におけるマッチングの結果をもとに仮想的な中間フレームを生成する工程と、

仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号する工程と、

を含むことを特徴とする画像データ復号方法。

【請求項13】 前記付加データは、前記現実の中間フレームと前記仮想的な中間フレームの差分をもとに生成されたデータである請求項12に記載の画像データ復号方法。

【請求項14】 前記復号する工程は、前記仮想的な中間フレームと前記差分をもとに生成されたデータの復号データを加算して前記現実の中間フレームを復号する請求項13に記載の画像データ復号方法。

【請求項15】 前記キーフレームのデータと前記現実の中間フレームのデータの組合せを前記画像データの復号データとして出力する工程をさらに含む請求項12から14のいずれかに記載の画像データ復号方法。

【請求項16】 コンピュータにて実行可能なプログラムを格納した記録媒体であって、前記プログラムは、

画像データの符号化データに含まれる前記画像データのキーフレームとその他の付加データを分別する工程と、

分別によって得られたキーフレームの間におけるマッチングの結果をもとに仮想的な中間フレームを生成する工程と、

仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号する工程と、

を含むことを特徴とする、コンピュータにて読み取り可能な記録媒体。

【請求項17】 画像データの符号化データを取得するユニットと、前記符号化データに含まれる前記画像データのキーフレームとその他の付加データを分別するユニットと、

分別によって得られたキーフレームの間でマッチングを計算するユニットと、マッチングの結果をもとに仮想的な中間フレームを生成するユニットと、仮想的な中間フレームと前記その他のデータを利用して現実の中間フレームを

復号するユニットと、

を含むことを特徴とする画像データ復号装置。

【請求項18】 前記付加データは、前記現実の中間フレームと前記仮想的な中間フレームの差分をもとに生成されたデータである請求項17に記載の画像データ復号装置。

【請求項19】 前記復号するユニットは、前記仮想的な中間フレームと前記差分をもとに生成されたデータの復号データを加算して前記現実の中間フレームを復号する請求項18に記載の画像データ復号装置。

【請求項20】 前記キーフレームのデータと前記現実の中間フレームのデータの組合せを前記画像データの復号データとして出力するユニットをさらに含む請求項17から19のいずれかに記載の画像データ復号装置。

【請求項21】 画像データを識別するためのインデックス領域と、
画像データに含まれるキーフレームのデータを格納するキーフレーム領域と、
前記キーフレーム領域に格納されたキーフレーム間のマッチングの計算結果を
もとに生成される中間フレームに関連するデータを格納する中間フレーム領域と

を含み、それらの領域のデータが前記画像データの符号化データとして関連づけられていることを特徴とする、コンピュータにて読み取り可能なデータ構造。

【請求項22】 前記中間フレーム領域は、前記キーフレーム間のマッチングの計算結果をもとに生成される仮想的な中間フレームを利用して符号化された、前記画像データに含まれる現実の中間フレームのデータを格納する請求項21に記載のデータ構造。

【請求項23】 画像データに含まれるフレームをキーフレームと中間フレームに分別する工程と、

分別によって得られた第1のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の始点階層画像を生成する工程と、

分別によって得られた第2のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の終点階層画像を生成する工程と、

始点階層画像および終点階層画像のマッチングを解像度レベルの階層の中で計

算する工程と、

マッチングの結果をもとに仮想的な中間フレームを生成する工程と、

仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する工程と、

を含むことを特徴とする画像データ符号化方法。

【請求項24】 複数のフレームを含む画像データを取得するユニットと、
前記画像データに含まれるフレームをキーフレームと中間フレームに分別する
ユニットと、

分別によって得られたキーフレームを入力し、それらキーフレーム間のマッチ
ングを計算するユニットと、

マッチングの計算結果をもとに仮想的な中間フレームを生成するユニットと、
仮想的な中間フレームを利用して前記分別によって得られた現実の中間フレ
ームを符号化するユニットとを含み、

前記マッチングを計算するユニットは、

分別によって得られた第1のキーフレームに多重解像度特異点フィルタを作用
させて解像度の異なる一連の始点階層画像を生成し、分別によって得られた第2
のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の
終点階層画像を生成し、始点階層画像および終点階層画像のマッチングを解像度
レベルの階層の中で計算することを特徴とする画像データ符号化装置。

【請求項25】 画像データに含まれるキーフレーム間でなされた処理の結
果をもとに生成された仮想的な中間フレームを取得し、

前記仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フ
ームを符号化することを特徴とする画像データ符号化方法。

【請求項26】 画像データに含まれるキーフレーム間でなされた処理の結
果をもとに生成された仮想的な中間フレームを取得する機能ブロックと、

前記仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フ
ームを符号化する機能ブロックと、

を含むことを特徴とする画像データ符号化装置。

【請求項27】 画像データの符号化データに含まれるキーフレームと付加

データの分別によって得られたキーフレーム間でなされた処理の結果から生成された仮想的な中間フレームを取得し、

仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号することを特徴とする画像データ復号方法。

【請求項28】 画像データの符号化データに含まれるキーフレームと付加データの分別によって得られたキーフレーム間でなされた処理の結果から生成された仮想的な中間フレームを取得する機能ブロックと、

仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号する機能ブロックと、

を含むことを特徴とする画像データ復号装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、画像データ処理技術に関する。この発明はとくに、複数のフレームを含む画像データを符号化または復号する方法と装置に関する。

【0002】

【従来の技術】

MRIやCTなどの医療画像装置の性能が向上し、従来と同じX線被曝量で非常に多くの断面画像が撮影できるようになった。たとえば従来せいぜい20枚しか撮影できなかった医療患部について、ここ数年の改良により、300枚や1000枚の撮影すら可能になっている。

【0003】

一方、それだけの枚数が撮影できても、これを法律の定めにしたがって所定期間保存する必要があり、データ量の増大が医療機関の悩みの種であった。また、病院内、病院間などで医療画像の伝送が必要な場合も、撮影画像の枚数がそのまま通信時間の増大を招き、場合により、医療の円滑化に逆行しかねない状況を生んでいた。

【0004】

【発明が解決しようとする課題】

こうした現状に対応すべく、厚生省は最近医療画像のデジタルデータ形式による保存を認めるようになった。デジタル化により、いろいろなデータ圧縮技術の活用が可能になるためである。しかし、取得される精緻の画像のデータ量増大に対して圧縮技術が追いつかないことは、ひとり医療分野に限らず、画像処理全般の状況である。したがって、さらに効率的な圧縮手法の探求は永遠のテーマといつてもよく、こうした技術に対する社会的要望は非常に強い。

【0005】

本発明はこうした現状に鑑みてなされたものであり、その目的は、画像データの効率的圧縮を実現する符号化および復号技術の提供にある。

【0006】

別の目的は、符号化の際に容易に画質の維持を考慮可能な画像処理技術の提供にある。

【0007】

また別の目的は、画質の維持と圧縮率の向上という相反する要望に適う画像符号化および復号技術の提供にある。

【0008】

【課題を解決するための手段】

本発明の画像データ符号化方法は、画像データに含まれるフレームをキーフレームと中間フレームに分別する工程と、分別によって得られたキーフレームの間でマッチングを計算する工程と、マッチングの結果をもとに仮想的な中間フレームを生成する工程と、仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する工程とを含む。

【0009】

「分別」は当初キーフレームと中間フレームの区別のなかったものを分類するという積極的な意味のほか、当初より区別がされていたものをその表示にしたがって分類するという消極的な意味を含む。

【0010】

「キーフレーム」はマッチングの対象になる参照フレームをいい、「中間フレーム」は対象にならない非参照フレームをいう。本明細書ではとくに断らない限

り、記述の簡潔のために、「フレーム」ということばを、画像の単位として呼ぶ場合と、その単位を構成するデータそのもの、すなわち「フレームデータ」と呼ぶべき場合の両義に用いる。

【0011】

「仮想的な中間フレーム」は、マッチング計算から導出されたフレームをいい、当初の画像データの中で「中間フレーム」と分別されたもの、すなわち「現実の中間フレーム」とは別個である。

【0012】

本発明で処理される画像データは、動画でもよいし、医療画像のように三次元オブジェクトを二次元化した静止画でもよい。さらにいえば、フレームとして把握できる任意次元の画像すべてが処理可能である。

【0013】

この態様によれば、現実の中間フレームが仮想的な中間フレームをもとに符号化される。一般則として、仮想的な中間フレームをそれと現実の中間フレームとの差分が少ない状態で生成できれば、その差分を圧縮符号化することで現実の中間フレームの符号量が少なくなる。仮想的な中間フレーム自体はマッチング計算から求まるため、それに固有の符号量は少なく、またはゼロにすることができる

【0014】

前記計算する工程は、キーフレーム間で画素単位のマッチングを計算し、前記生成する工程は、キーフレーム間における画素の位置および輝度の対応関係をもとに、画素単位で補間計算を実施して仮想的な中間フレームを生成してもよい。「補間」は内挿補間のほか、外挿補間でもよいし、線形、非線形いずれでもよい

【0015】

本発明の画像データ符号化方法はさらに、キーフレームのデータと前記符号化する工程で得られたデータの組合せを画像データの符号化データとして出力、保存または伝送する工程を含んでもよい。したがって、本明細書で「符号化」というと、中間フレームに関するものと、画像データ全体に関するものがあるが、一

般に前者の結果後者が生じる。

【0016】

本発明の画像データ符号化装置は、複数のフレームを含む画像データを取得するユニットと、前記画像データに含まれるフレームをキーフレームと中間フレームに分別するユニットと、分別によって得られたキーフレームを入力し、それらキーフレーム間のマッチングを計算するユニットと、マッチングの計算結果をもとに仮想的な中間フレームを生成するユニットと、仮想的な中間フレームを利用して前記分別によって得られた現実の中間フレームを符号化するユニットとを含む。これらの各ユニットは、ソフトウェア、ハードウェアの任意の組合せで実現できる。

【0017】

前記生成するユニットは、前記画素単位のマッチングの結果をもとに前記キーフレームの画素間を補間することで前記仮想的な中間フレームを生成してもよい。

【0018】

前記生成するユニットはまた、前記キーフレーム間における画素の位置および輝度の対応関係をもとに、画素単位で補間計算を実施して前記仮想的な中間フレームを生成してもよい。

【0019】

一方、本発明の画像データ復号方法は、画像データの符号化データに含まれる前記画像データのキーフレームとその他の付加データを分別する工程と、分別によって得られたキーフレームの間におけるマッチングの結果をもとに仮想的な中間フレームを生成する工程と、仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号する工程とを含む。本明細書において、「復号」は中間フレームに関するものと画像データ全体に関するものがあるが、一般に前者の結果後者が生じる。

【0020】

前記付加データは、前記現実の中間フレームと前記仮想的な中間フレームの差分をもとに生成されたデータ、たとえば差分にエントロピー符号化、JPEGな

どの空間周波数に関する符号化を施して得られたデータであってもよい。その場合、前記復号する工程は、前記仮想的な中間フレームと前記差分をもとに生成されたデータの復号データを加算して前記現実の中間フレームを復号してもよい。

【0021】

本発明の画像データ復号方法はさらに、前記キーフレームのデータと前記現実の中間フレームのデータの組合せを前記画像データの復号データとして出力、保存、伝送する工程を含んでもよい。

【0022】

本発明の画像データ復号装置は、画像データの符号化データを取得するユニットと、前記符号化データに含まれる前記画像データのキーフレームとその他の付加データを分別するユニットと、分別によって得られたキーフレームの間でマッチングを計算するユニットと、マッチングの結果をもとに仮想的な中間フレームを生成するユニットと、仮想的な中間フレームと前記その他のデータを利用して現実の中間フレームを復号するユニットとを含む。

【0023】

本発明の画像データ符号化方法のさらに別の態様は、画像データに含まれるフレームをキーフレームと中間フレームに分別する工程と、分別によって得られた第1のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の始点階層画像を生成する工程と、分別によって得られた第2のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の終点階層画像を生成する工程と、始点階層画像および終点階層画像のマッチングを解像度レベルの階層の中で計算する工程と、マッチングの結果をもとに仮想的な中間フレームを生成する工程と、仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する工程とを含む。

【0024】

本発明の画像データ符号化装置のさらに別の態様は、複数のフレームを含む画像データを取得するユニットと、前記画像データに含まれるフレームをキーフレームと中間フレームに分別するユニットと、分別によって得られたキーフレームを入力し、それらキーフレーム間のマッチングを計算するユニットと、マッチ

グの計算結果をもとに仮想的な中間フレームを生成するユニットと、仮想的な中間フレームを利用して前記分別によって得られた現実の中間フレームを符号化するユニットとを含む。

【0025】

この構成で、前記マッチングを計算するユニットは、分別によって得られた第1のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の始点階層画像を生成し、分別によって得られた第2のキーフレームに多重解像度特異点フィルタを作用させて解像度の異なる一連の終点階層画像を生成し、始点階層画像および終点階層画像のマッチングを解像度レベルの階層の中で計算する。

【0026】

本発明の画像データ符号化方法の別の態様は以下の処理を含む。すなわち、画像データに含まれるキーフレーム間でなされた処理の結果をもとに生成された仮想的な中間フレームを取得し、前記仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する。つまり、マッチングまたはそれと仮想的な中間フレームの生成等の処理を本発明の前処理と捉えるものである。

【0027】

同様の趣旨で、本発明の画像データ符号化装置の別の態様は、画像データに含まれるキーフレーム間でなされた処理の結果をもとに生成された仮想的な中間フレームを取得する機能ブロックと、前記仮想的な中間フレームを利用して前記画像データに含まれる現実の中間フレームを符号化する機能ブロックとを含む。

【0028】

本発明の画像データ復号方法の別の態様は、画像データの符号化データに含まれるキーフレームと付加データの分別によって得られたキーフレーム間でなされた処理の結果から生成された仮想的な中間フレームを取得し、仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号する。仮想的な中間フレームとキーフレームの入力から処理を開始する趣旨である。

【0029】

同様の趣旨で、本発明の画像データ復号装置の別の態様は、画像データの符号化データに含まれるキーフレームと付加データの分別によって得られたキーフレーム間でなされた処理の結果から生成された仮想的な中間フレームを取得する機能ブロックと、仮想的な中間フレームと前記付加データを利用して現実の中間フレームを復号する機能ブロックとを含む。

【0030】

【発明の実施の形態】

はじめに、実施の形態で利用する多重解像度特異点フィルタ技術とそれを用いた画像マッチング処理を「前提技術」として詳述する。これらの技術は本出願人がすでに特許第2927350号を得ている技術であり、本発明との組合せに適する。ただし、実施の形態で採用可能な画像マッチング技術はこれに限られない。

【0031】

図23以降、前提技術を利用した画像データ符号化および復号技術を具体的に説明する。

[前提技術に関する従来の技術]

ふたつの画像の自動的なマッチング、つまり画像領域や画素どうしの対応付けは、コンピュータビジョンやコンピュータグラフィックスにおける最も難しくかつ重要なテーマのひとつである。例えば、あるオブジェクトに関して異なる視点からの画像間でマッチングがとれれば、他の視点からの画像を生成することができる。右目画像と左目画像のマッチングが計算できれば、立体画像を用いた写真測量も可能である。顔の画像のモデルと他の顔の画像のマッチングがとれたとき、目、鼻、口といった特徴的な顔の部分を抽出することができる。例えば人の顔と猫の顔の画像間でマッチングが正確にとられたとき、それらの中割画像を自動的に生成することでモーフィングを完全自動化することができる。

【0032】

しかし従来一般に、ふたつの画像間の対応点は人がいちいち指定しなければならず、多大な作業工数を要した。この問題を解消するために数多くの対応点自動検出方法が提案されている。例えば、エピポーラ直線を用いることによって対応

点の候補の数を減らす考えがある。しかし、その場合でも処理はきわめて複雑である。複雑さを低減するために、左目画像の各点の座標は通常右目画像でもほぼ同じ位置にあると想定される。しかし、こうした制約を設けると、大域的特徴及び局所的特徴を同時に満たすマッチングをとることは非常に困難になる。

【0033】

ボリュームレンダリングでは、ボクセルを構成するために一連の断面画像が用いられる。この場合、従来一般に、上方の断面画像における画素が下方の断面画像の同一箇所にある画素と対応すると仮定され、これらの画素のペアが内挿計算に用いられる。このようにきわめて単純な方法を用いるため、連続する断面間の距離が遠く、オブジェクトの断面形状が大きく変化する場合、ボリュームレンダリングで構築されたオブジェクトは不明瞭になりがちである。

【0034】

立体写真測量法など、エッジの検出を利用するマッチングアルゴリズムも多い。しかしこの場合、結果的に得られる対応点の数が少ないため、マッチングのとれた対応点間のギャップを埋めるべく、ディスパリティの値を内挿計算しなければならない。一般にあらゆるエッジ検出器は、それらが用いる局所的なウインドウの中で画素の輝度が変化したとき、これが本当にエッジの存在を示唆するかどうかを判断することが難しい。エッジ検出器は、本来的にすべてハイパスフィルタであり、エッジと同時にノイズも拾ってしまう。

【0035】

さらに別の手法として、オプティカルフローが知られている。二枚の画像が与えられたとき、オプティカルフローでは画像内のオブジェクト（剛体）の動きを検出する。その際、オブジェクトの各画素の輝度は変化しないと仮定する。オプティカルフローでは例えば (u, v) のベクトル場の滑らかさといった、いくつかの付加的な条件とともに、各画素の動きベクトル (u, v) を計算する。しかし、オプティカルフローでは画像間の大域的な対応関係を検出することはできない。画素の輝度の局所的な変化に注目するのみであり、画像の変位が大きい場合、システムの誤差は顕著になる。

【0036】

画像の大域的な構造を認識するために、多重解像度フィルタも数多く提案されてきた。それらは線形フィルタと非線形フィルタに分類される。前者の例としてウェーブレットがあるが、線形フィルタは一般に、画像マッチングにはさして有用ではない。なぜなら、極値をとる画素の輝度に関する情報がそれらの位置情報とともに次第に不鮮明になるためである。図1 (a) と図1 (b) はそれぞれ図19 (a) と図19 (b) に示す顔の画像に対して平均化フィルタを適用した結果を示している。同図のごとく、極値をとる画素の輝度が平均化によって次第に薄れるとともに、位置も平均化の影響でシフトしていく。その結果、目（輝度の極小点）の輝度や位置の情報は、このような粗い解像度レベルで曖昧になり、この解像度では正しいマッチングを計算することができない。したがって、粗い解像度レベルを設けるのが大域的なマッチングのためでありながら、ここで得られたマッチングは画像の本当の特徴（目、つまり極小点）に正確に対応しない。より精細な解像度レベルで目が鮮明に現れたとしても、大域的なマッチングをとる際に混入した誤差は、もはや取り返しがつかない。入力画像にスムージング処理を加えることにより、テクスチャ領域のステレオ情報が落ちてしまうこともすでに指摘されている。

【0037】

一方、最近地形学の分野で利用されはじめた非線形フィルタとして一次元の「ふるい (sieve)」演算子がある。この演算子は、所定の大きさの一次元ウィンドウ内の極小値（または極大値）を選択することにより、縮尺と空間の因果関係を保存しながら画像にスムージング処理を加える。その結果得られる画像は元の画像と同じ大きさであるが、小さな波の成分が取り除かれているため、より単純になる。画像の情報を落とすという点で、この演算子は広い意味での「多重解像度フィルタ」に分類することはできるが、実際にはウェーブレットのように画像の解像度を変えながら画像を階層化するわけではなく（つまり狭い意味での多重解像度フィルタではなく）、画像間の対応の検出には利用できない。

[前提技術が解決しようとする課題]

以上をまとめれば以下の課題が認められる。

【0038】

1. 画像の特徴を正確に、かつ比較的簡単な処理で把握する画像処理方法が乏しかった。特に、特徴のある点に関する情報、例えば画素値や位置を維持しながら特徴を抽出できる画像処理方法に関する有効な提案が少なかった。

2. 画像の特徴をもとに対応点を自動検出する場合、一般に処理が複雑であるか、ノイズ耐性が低いなどの欠点があった。また、処理に際していろいろな制約を設ける必要があり、大域的特徴及び局所的特徴を同時に満たすマッチングをとることが困難だった。

3. 画像の大域的な構造または特徴を認識するために多重解像度フィルタを導入しても、そのフィルタが線形フィルタの場合、画素の輝度情報と位置情報が曖昧になった。その結果、対応点の把握が不正確になりやすかった。非線形フィルタである一次元ふるい演算子は画像を階層化しないため、画像間の対応点の検出には利用できなかった。

4. これらの結果、対応点を正しく把握しようとすれば、結局人手による指定に頼るほか有効な手立てがなかった。

【0039】

本発明はこれらの課題の解決を目的としてなされたものであり、画像処理の分野において、画像の特徴の的確な把握を可能にする技術を提供するものである。

[前提技術が課題を解決するための手段]

この目的のために本発明のある態様は、新たな多重解像度の画像フィルタを提案する。この多重解像度フィルタは画像から特異点を抽出する。したがって、特異点フィルタともよばれる。特異点とは画像上特徴をもつ点をいう。例として、ある領域において画素値（画素値とは、色番号、輝度値など画像または画素に関する任意の数値を指す）が最大になる極大点、最小になる極小点、ある方向については最大だが別の方向については最小になるような鞍点がある。特異点は位相幾何学上の概念であってもよい。ただし、その他どのような特徴を有してもよい。いかなる性質の点を特異点と考えるかは、本発明にとって本質問題ではない。

【0040】

この態様では、多重解像度フィルタを用いた画像処理が行われる。まず検出工程において、第一の画像に対し、二次元的な探索を行って特異点が検出される。

つぎに生成工程において、検出された特異点を抽出して第一の画像よりも解像度の低い第二の画像が生成される。第二の画像には第一の画像のもつ特異点が引き継がれる。第二の画像は第一の画像よりも解像度が低いため、画像の大域的な特徴の把握に好適である。

【0041】

本発明の別の態様は特異点フィルタを用いた画像マッチング方法に関する。この態様では、始点画像と終点画像間のマッチングがとられる。始点画像および終点画像とは、ふたつの画像の区別のために便宜的に与えた名称であり、本質的な違いはない。

【0042】

この態様では、第一工程にて、始点画像に特異点フィルタを施して解像度の異なる一連の始点階層画像が生成される。第二工程では、終点画像に特異点フィルタを施して解像度の異なる一連の終点階層画像が生成される。始点階層画像、終点階層画像とは、それぞれ始点画像、終点画像を階層化して得られる画像群をいい、それぞれ最低2枚の画像からなる。つぎに第三工程において、始点階層画像と終点階層画像のマッチングが解像度レベルの階層の中で計算される。この態様によれば、多重解像度フィルタによって特異点に関連する画像の特徴が抽出され、および／または明確化されるため、マッチングが容易になる。マッチングのための拘束条件は特に必要としない。

【0043】

本発明のさらに別の態様も始点画像と終点画像のマッチングに関する。この態様では、予め複数のマッチング評価項目のそれぞれに関して評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索する。総合評価式は、評価式の少なくもひとつに係数パラメータを掛けたうえでそれらの評価式の総和として定義してもよく、その場合、総合評価式またはいずれかの評価式がほぼ極値をとる状態を検出して前記パラメータを決定してもよい。「極値付近」または「ほぼ極値をとる」としたのは、多少誤差を含んでいてもよいためである。多少の誤差は本発明にはさして問題とならない。

【0044】

極値自体も前記パラメータに依存するため、極値の挙動、つまり極値の変化の様子をもとに、最適と考えられるパラメータを決定する余地が生じる。この態様はその事実を利用している。この態様によれば、元来調整の困難なパラメータの決定を自動化する途が拓かれる。

[前提技術の実施の形態]

最初に〔1〕で前提技術の要素技術の詳述し、〔2〕で処理手順を具体的に説明する。さらに〔3〕で実験の結果を報告する。

〔1〕要素技術の詳細

〔1. 1〕イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから精細なレベルへと順に解像度の階層を辿っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

【0045】

本前提技術は、例えば完全に自動的なモーフィング、物体認識、立体写真測量、ボリュームレンダリング、少ないフレームからの滑らかな動画像の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ボリュームレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

【0046】

〔1. 2〕特異点フィルタの階層

前提技術に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の輝度及び位置を保存することができる。ここで画像の幅をN、高さをMとする。以下簡単のため、 $N = M = 2^n$ (n は自然数)と仮定する。また、区間 $[0, N] \subset \mathbb{R}$ をIと記述する。 (i, j) における画

像の画素を $p(i, j)$ と記述する ($i, j \in I$)。

【0047】

ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第 m レベルにおける各画像のサイズは $2^m \times 2^m$ ($0 \leq m \leq n$) とする。特異点フィルタは次の4種類の新たな階層画像を n から下がる方向で再帰的に構築する。

【0048】

【数1】

$$\begin{aligned} p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\ p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\ p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\ p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)})) \end{aligned}$$

(式1)

ただしここで、

【数2】

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式2})$$

とする。以降これら4つの画像を副画像(サブイメージ)と呼ぶ。 $\min_{x \leq t} \leq x+1$ 、 $\max_{x \leq t} x \leq x+1$ をそれぞれ α 及び β と記述すると、副画像はそれぞれ以下のように記述できる。

【0049】

$$\begin{aligned} P(m, 0) &= \alpha(x) \alpha(y) p(m+1, 0) \\ P(m, 1) &= \alpha(x) \beta(y) p(m+1, 1) \\ P(m, 2) &= \beta(x) \alpha(y) p(m+1, 2) \\ P(m, 3) &= \beta(x) \beta(y) p(m+1, 3) \end{aligned}$$

すなわち、これらは α と β のテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明らかのように、特異点フィルタはもとの画像について 2×2 画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、前提技術では輝度を採用するが、画像に関するいろいろな数値を採用することができる。ふたつの方向の両方について最大画素値となる画素は極大点、ふたつの方向の両方について最小画素値となる画素は極小点、ふたつの方向の一方について最大画素値となるとともに、他方について最小画素値となる画素は鞍点として検出される。

【0050】

特異点フィルタは、各ブロックの内部で検出された特異点の画像（ここでは1画素）でそのブロックの画像（ここでは4画素）を代表させることにより、画像の解像度を落とす。特異点の理論的な観点からすれば、 $\alpha(x)\alpha(y)$ は極小点を保存し、 $\beta(x)\beta(y)$ は極大点を保存し、 $\alpha(x)\beta(y)$ 及び $\beta(x)\alpha(y)$ は鞍点を保存する。

【0051】

はじめに、マッチングをとるべき始点（ソース）画像と終点（デスティネーション）画像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

【0052】

この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとれらていく。まず $p(m, 0)$ を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p(m, 1)$ を用いて鞍点のマッチングがとられ、 $p(m, 2)$ を用いて他の鞍点のマッチングがとられる。そして最後に $p(m, 3)$ を用いて極大点のマッチングがとられる。

【0053】

図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像 $p(5, 0)$ を示している。同様に、図1(e)と図1(f)は $p(5, 1)$ 、図1(g)

) と図1 (h) は $p(5, 2)$ 、図1 (i) と図1 (j) は $p(5, 3)$ をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず $p(5, 0)$ によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p(5, 1)$ によれば口が明確になる。口は横方向で輝度が低いためである。 $p(5, 2)$ によれば首の両側の縦線が明確になる。最後に、 $p(5, 3)$ によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

【0054】

特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影された画像の特徴と、予め記録しておいたいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

【0055】

[1. 3] 画像間の写像の計算

始点画像の位置 (i, j) の画素を $p(n)(i, j)$ と書き、同じく終点画像の位置 (k, l) の画素を $q(n)(k, l)$ で記述する。 $i, j, k, l \in I$ とする。画像間の写像のエネルギー（後述）を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ $p(m, 0)$ と $q(m, 0)$ 間の写像 $f(m, 0) : p(m, 0) \rightarrow q(m, 0)$ が計算される。 $f(m, 0)$ に基づき、最小エネルギーを持つ $p(m, 1), q(m, 1)$ 間の写像 $f(m, 1)$ が計算される。この手続は、 $p(m, 3)$ と $q(m, 3)$ の間の写像 $f(m, 3)$ の計算が終了するまで続く。各写像 $f(m, i)$ ($i = 0, 1, 2, \dots$) を副写像と呼ぶことにする。 $f(m, i)$ の計算の都合のために、 i の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

【数3】

$$f^{(m,i)} : p^{(m,\sigma(i))} \rightarrow q^{(m,\sigma(i))} \quad (\text{式3})$$

ここで $\sigma(i) \in \{0, 1, 2, 3\}$ である。

【0056】

[1. 3. 1] 全单射

始点画像と終点画像の間のマッチングを写像で表現する場合、その写像は両画像間で全单射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射かつ单射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全单射のディジタル版である。前提技術では、画素は格子点によって特定される。

【0057】

始点副画像（始点画像について設けられた副画像）から終点副画像（終点画像について設けられた副画像）への写像は、 $f(m, s) : I/2n - m \times I/2n - m \rightarrow I/2n - m \times I/2n - m$ ($s = 0, 1, \dots$) によって表される。ここで、 $f(m, s)(i, j) = (k, l)$ は、始点画像の $p(m, s)(i, j)$ が終点画像の $q(m, s)(k, l)$ に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$ が成り立つとき画素 $q(k, l)$ を $q_f(i, j)$ と記述する。

【0058】

前提技術で扱う画素（格子点）のようにデータが離散的な場合、全单射の定義は重要である。ここでは以下のように定義する（ i, i', j, j', k, l は全て整数とする）。まず始めに、始点画像の平面において R によって表記される各正方形領域、

【数4】

$$p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}$$

(式4)

を考える ($i = 0, \dots, 2m - 1, j = 0, \dots, 2m - 1$)。ここで R の各辺（エッジ）の方向を以下のように定める。

【0059】

【数5】

$$\overrightarrow{p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)}}, \overrightarrow{p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)}}, \overrightarrow{p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{p_{(i,j+1)}^{(m,s)} p_{(i,j)}^{(m,s)}}$$

(式5)

この正方形は写像 f によって終点画像平面における四辺形に写像されなければ

ならない。 $f(m, s)(R)$ によって示される四辺形、

【数6】

$$q_{f(i,j)}^{(m,s)} q_{f(i+1,j)}^{(m,s)} q_{f(i+1,j+1)}^{(m,s)} q_{f(i+1,j+1)}^{(m,s)} \quad (\text{式6})$$

は、以下の全单射条件を満たす必要がある。

【0060】

1. 四辺形 $f(m, s)(R)$ のエッジは互いに交差しない。
2. $f(m, s)(R)$ のエッジの方向は R のそれらに等しい（図2の場合、時計回り）。
3. 緩和条件として収縮写像（リトラクション：retractions）を許す。

【0061】

何らかの緩和条件を設けないかぎり、全单射条件を完全に満たす写像は単位写像しかないと認められる。ここでは $f(m, s)(R)$ の一つのエッジの長さが 0、すなわち $f(m, s)(R)$ は三角形になってもよい。しかし、面積が 0 となるような図形、すなわち 1 点または 1 本の線分になってはならない。図2 (R) がもとの四辺形の場合、図2 (A) と図2 (D) は全单射条件を満たすが、図2 (B)、図2 (C)、図2 (E) は満たさない。

【0062】

実際のインプリメンテーションでは、写像が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというものである。すなわち、 $f(i, j) = (i, j)$ （ただし $i = 0, i = 2m - 1, j = 0, j = 2m - 1$ の 4 本の線上）である。この条件を以下「付加条件」とも呼ぶ。

【0063】

[1. 3. 2] 写像のエネルギー

[1. 3. 2. 1] 画素の輝度に関するコスト

写像 f のエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像 $f(m, s)$ の点 (i, j) における

るエネルギーC (m, s) (i, j) は次式によって定まる。

【0064】

【数7】

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_{f(i,j)}^{(m,s)})|^2 \quad (\text{式7})$$

ここで、V (p (m, s) (i, j)) 及びV (q (m, s) f (i, j)) はそれぞれ画素p (m, s) (i, j) 及びq (m, s) f (i, j) の輝度である。f のトータルのエネルギーC (m, s) は、マッチングを評価するひとつの評価式であり、つぎに示すC (m, s) (i, j) の合計で定義できる。

【0065】

【数8】

$$C_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} C_{(i,j)}^{(m,s)} \quad (\text{式8})$$

【0066】

【1. 3. 2. 2】滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギーD f を導入する。このエネルギーは画素の輝度とは関係なく、p (m, s) (i, j) およびq (m, s) f (i, j) の位置によって決まる (i = 0, …, 2m - 1, j = 0, …, 2m - 1)。点 (i, j) における写像 f (m, s) のエネルギーD (m, s) (i, j) は次式で定義される。

【0067】

【数9】

$$D_{(i,j)}^{(m,s)} = \eta E_{0(i,j)}^{(m,s)} + E_{1(i,j)}^{(m,s)} \quad (\text{式9})$$

ただし、係数パラメータ η は0以上の実数であり、また、

【数10】

$$E_{0(i,j)}^{(m,s)} = \|(i, j) - f^{(m,s)}(i, j)\|^2 \quad (\text{式10})$$

【数11】

$$E_{1(i,j)}^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2/4 \quad (式11)$$

とする。ここで、

【数12】

$$\|(x,y)\| = \sqrt{x^2 + y^2}$$

であり、 $i' < 0$ および $j' < 0$ に対して $f^{(i', j')}$ は 0 と決める。 E_0 は (i, j) 及び $f^{(i, j)}$ の距離で決まる。 E_0 は画素があまりにも離れた画素へ写影されることを防ぐ。ただし E_0 は、後に別のエネルギー関数で置き換える。 E_1 は写像の滑らかさを保証する。 E_1 は、 $p(i, j)$ の変位とその隣接点の変位の間の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー D_f は次式で定まる。

【0068】

【数13】

$$D_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} D_{(i,j)}^{(m,s)} \quad (式13)$$

[1. 3. 2. 3] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は $\lambda C^{(m,s)}_f + D^{(m,s)}_f$ で定義される。ここで係数パラメータ λ は 0 以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次式で示す最小エネルギーを与える写像を見いだすことである。

【数14】

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \quad (式14)$$

【0069】

$\lambda = 0$ 及び $\eta = 0$ の場合、写像は単位写像になることに注意すべきである（すなわち、全ての $i = 0, \dots, 2^m - 1$ 及び $j = 0, \dots, 2^m - 1$ に対して $f^{(m,s)}(i, j) = (i, j)$ となる）。後述のごとく、本前提技術では最初に

$\lambda = 0$ 及び $\eta = 0$ の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の λ の位置を変えて $C(m, s) f + \lambda D(m, s) f$ と定義したとすれば、 $\lambda = 0$ 及び $\eta = 0$ の場合に総合評価式が $C(m, s) f$ だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけで対応づけられ、写像が無意味なものになる。そうした無意味な写像をもとに写像を変形していってもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

【0070】

オプティカルフローもこの前提技術同様、画素の輝度の差と滑らかさを考慮する。しかし、オプティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。前提技術に係る特異点フィルタを用いることによって大域的な対応関係を検出することができる。

【0071】

[1. 3. 3] 多重解像度の導入による写像の決定

最小エネルギーを与え、全单射条件を満足する写像 f_{min} を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0072】

まず、

【数15】

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式15})$$

が成り立つとき、 $p(m-1, s)(i', j')$ 、 $q(m-1, s)(i', j')$ をそれぞれ $p(m, s)(i, j)$ 、 $q(m, s)(i, j)$ のpare

nt と呼ぶことにする。 $[x]$ は x を越えない最大整数である。また $p(m, s)$ (i, j)、 $q(m, s)$ (i, j) をそれぞれ $p(m-1, s)$ (i', j')、 $q(m-1, s)$ (i', j') の child と呼ぶ。関数 `parent` (i, j) は次式で定義される。

【0073】

【数16】

$$parent(i, j) = \left(\left[\frac{i}{2} \right], \left[\frac{j}{2} \right] \right) \quad (式16)$$

$p(m, s)$ (i, j) と $q(m, s)$ (k, l) の間の写像 $f(m, s)$ は、エネルギー計算を行って最小になったものを見つけることで決定される。 $f(m, s)$ (i, j) = (k, l) の値は $f(m-1, s)$ ($m=1, 2, \dots, n$) を用いることによって、以下のように決定される。まず、 $q(m, s)$ (k, l) は次の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

【0074】

【数17】

$$q_{g(m,s)}^{(m,s)}(i-1, j-1) q_{g(m,s)}^{(m,s)}(i-1, j+1) q_{g(m,s)}^{(m,s)}(i+1, j+1) q_{g(m,s)}^{(m,s)}(i+1, j-1)$$

(式17)

ただしここで、

【数18】

$$g^{(m,s)}(i, j) = f^{(m-1,s)}(parent(i, j)) + f^{(m-1,s)}(parent(i, j) + (1, 1))$$

(式18)

である。こうして定めた四辺形を、以下 $p(m, s)$ (i, j) の相続 (inherited) 四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0075】

図3は以上の手順を示している。同図において、始点画像の A, B, C, D の画素は、第 $m-1$ レベルにおいてそれぞれ終点画像の A', B', C', D' へ

写影される。画素 $p (m, s) (i, j)$ は、相続四辺形 $A' B' C' D'$ の内部に存在する画素 $q (m, s) f (m) (i, j)$ へ写影されなければならない。以上の配慮により、第 $m-1$ レベルの写像から第 m レベルの写像への橋渡しがなされる。

【0076】

先に定義したエネルギー E_0 は、第 m レベルにおける副写像 $f (m, 0)$ を計算するために、次式に置き換える。

【0077】

【数19】

$$E_{0(i,j)} = \|f^{(m,0)}(i,j) - g^{(m)}(i,j)\|^2 \quad (\text{式19})$$

また、副写像 $f (m, s)$ を計算するためには次式を用いる。

【数20】

$$E_{0(i,j)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2 \quad (1 \leq s) \quad (\text{式20})$$

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式20により、異なる特異点に対応する副写像が、副写像どうしの類似度が高くなるように同一レベル内で関連づけられる。式19は、 $f (m, s) (i, j)$ と、第 $m-1$ レベルの画素の一部と考えた場合の (i, j) が射影されるべき点の位置との距離を示している。

【0078】

仮に、相続四辺形 $A' B' C' D'$ の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、 $A' B' C' D'$ の境界線からの距離が L (始めは $L = 1$) である画素を調べる。それらのうち、エネルギーが最小になるものが全単射条件を満たせば、これを $f (m, s) (i, j)$ の値として選択する。そのような点が発見されるか、または L がその上限の $L (m) \max$ に到達するまで、 L を大きくしていく。 $L (m) \max$ は各レベル m に対して固定である。そのような点が全く発見されない場合、全単射の第3の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f (m, s) (i, j)$ を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の

第1及び第2条件を外す。

【0079】

多重解像度を用いる近似法は、写像が画像の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことができない。さらに、通常写像に滑らかさを要求するため、こうした画素間の対応関係を見つけにくくしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、こうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル（粗いレベル）において小さいためである。

【0080】

[1. 4] 最適なパラメータ値の自動決定

既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。前提技術に係る方法によれば、最適なパラメータ値を完全に自動決定することができる。

【0081】

前提技術に係るシステムはふたつのパラメータ、 λ 及び η を含む。端的にいえば、 λ は画素の輝度の差の重みであり、 η は写像の剛性を示している。これらのパラメータの値は初期値が 0 であり、まず $\eta = 0$ に固定して λ を 0 から徐々に増加させる。 λ の値を大きくしながら、しかも総合評価式（式 1.4）の値を最小にする場合、各副写像に関する $C(m, s) f$ の値は一般に小さくなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、 λ が最適値を超えると以下の現象が発生する。

【0082】

1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。
2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる

3. その結果、式14において $D(m, s) f$ が急激に増加しようとする。

4. その結果、式14の値が急激に増加しようとするため、 $D(m, s) f$ の急激な増加を抑制するよう $f(m, s)$ が変化し、その結果 $C(m, s) f$ が増加する。

したがって、 λ を増加させながら式14が最小値をとるという状態を維持しつつ $C(m, s) f$ が減少から増加に転じる閾値を検出し、その λ を $\eta = 0$ における最適値とする。つぎに η を少しづつ増やして $C(m, s) f$ の挙動を検査し、後述の方法で η を自動決定する。その η に対応して λ も決まる。

【0083】

この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できるとき、その目が固定される。

【0084】

【1. 4. 1】 λ の動的決定

λ は0から所定の刻み幅で増加されていき、 λ の値が変わる度に副写像が評価される。式14のごとく、総エネルギーは $\lambda C(m, s) f + D(m, s) f$ によって定義される。式9の $D(m, s) f$ は滑らかさを表すもので、理論的には単位写像の場合に最小になり、写像が歪むほど E_0 も E_1 も増加していく。 E_1 は整数であるから、 $D(m, s) f$ の最小刻み幅は1である。このため、現在の $\lambda C(m, s) (i, j)$ の変化(減少量)が1以上でなければ、写像を変化させることによって総エネルギーを減らすことはできない。なぜなら、写像の変化に伴って $D(m, s) f$ は1以上増加するため、 $\lambda C(m, s) (i, j)$ が1以上減少しない限り総エネルギーは減らないためである。

【0085】

この条件のもと、 λ の増加に伴い、正常な場合に $C(m, s) (i, j)$ が減少することを示す。 $C(m, s) (i, j)$ のヒストグラムを $h(1)$ と記述する。 $h(1)$ はエネルギー $C(m, s) (i, j)$ が12である画素の数である。 $\lambda 12 \geq 1$ が成り立つために、例えば $12 = 1 / \lambda$ の場合を考える。 λ が $\lambda 1$

から λ_2 まで微小量変化するとき、

【0086】

【数21】

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{l=\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式21)

で示されるA個の画素が、

【数22】

$$C_f^{(m,s)} - l^2 = C_f^{(m,s)} - \frac{1}{\lambda} \quad (式22)$$

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素のエネルギーがすべてゼロになると近似している。この式は $C_f^{(m,s)}$ の値が、

【0087】

【数23】

$$\partial C_f^{(m,s)} = -\frac{4}{\lambda} \quad (式23)$$

だけ変化することを示し、その結果、

【0088】

【数24】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (式24)$$

が成立する。 $h(l) > 0$ であるから、通常 $C_f^{(m,s)}$ は減少する。しかし、 λ が最適値を越えようとするとき、上述の現象、つまり $C_f^{(m,s)}$ の増加が発生する。この現象を検出することにより、 λ の最適値を決定する。

【0089】

なお、 $H (h > 0)$ 及び k を定数とするとき、

【数25】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (式25)$$

と仮定すれば、

【数26】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (\text{式26})$$

が成り立つ。このとき $k \neq -3$ であれば、

【0090】

【数27】

$$C_f^{(m,s)} = C + \frac{H}{(3/2+k/2)\lambda^{3/2+k/2}} \quad (\text{式27})$$

となる。これが $C(m, s) f$ の一般式である (C は定数)。

【0091】

λ の最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率を p_0 と仮定する。この場合、

【0092】

【数28】

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式28})$$

が成立しているため、全単射条件を破る画素の数は次式の率で増加する。

【数29】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}} \quad (\text{式29})$$

従って、

【数30】

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1 \quad (\text{式30})$$

は定数である。仮に $h(l) = H l k$ を仮定するとき、例えば、

【数31】

$$B_0 \lambda^{3/2+k/2} = p_0 H \quad (\text{式31})$$

は定数になる。しかし λ が最適値を越えると、上の値は急速に増加する。この現

象を検出し、 $B0\lambda 3/2 + k/2/2$ mの値が異常値B0thresを越えるかどうかを検査し、 λ の最適値を決定することができる。同様に、 $B1\lambda 3/2 + k/2/2$ mの値が異常値B1thresを越えるかどうかを検査することにより、全単射の第3の条件を破る画素の増加率B1を確認する。ファクター2mを導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギーC(m, s)fの観察では検出し損なった写像の過度の歪みを検出するために用いることができる。

【0093】

なお実験では、副写像f(m, s)を計算する際、もし λ が0.1を越えたらf(m, s)の計算は止めてf(m, s+1)の計算に移行した。 $\lambda > 0.1$ のとき、画素の輝度255レベル中のわずか「3」の違いが副写像の計算に影響したためであり、 $\lambda > 0.1$ のとき正しい結果を得ることは困難だったためである。

【0094】

[1. 4. 2] ヒストグラムh(1)

C(m, s)fの検査はヒストグラムh(1)に依存しない。全単射及びその第3の条件の検査の際、h(1)に影響を受けうる。実際に($\lambda, C(m, s)f$)をプロットすると、kは通常1付近にある。実験ではk=1を用い、 $B0\lambda 2$ と $B1\lambda 2$ を検査した。仮にkの本当の値が1未満であれば、 $B0\lambda 2$ と $B1\lambda 2$ は定数にならず、ファクター $\lambda(1-k)/2$ に従って徐々に増加する。h(1)が定数であれば、例えばファクターは $\lambda 1/2$ である。しかし、こうした差は閾値B0thresを正しく設定することによって吸収することができる。

【0095】

ここで次式のごとく始点画像を中心が(x0, y0)、半径rの円形のオブジェクトであると仮定する。

【数32】

$$p_{(i,j)} = \begin{cases} \frac{255}{r} c \left(\sqrt{(i-x_0)^2 + (j-y_0)^2} \right) & \left(\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r \right) \\ 0 & \left(\text{otherwise} \right) \end{cases} \quad (\text{式32})$$

一方、終点画像は、次式のごとく中心(x1, y1)、半径がrのオブジェクト

トであるとする。

【0096】

【数33】

$$q_{(i,j)} = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式33})$$

ここで $c(x)$ は $c(x) = x^k$ の形であるとする。中心 (x_0, y_0) 及び (x_1, y_1) が十分遠い場合、ヒストグラム $h(l)$ は次式の形となる。

【0097】

【数34】

$$h(l) \propto rl^k \quad (k \neq 0) \quad (\text{式34})$$

$k = 1$ のとき、画像は背景に埋め込まれた鮮明な境界線を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。 $k = -1$ のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従って暗くなる。一般的のオブジェクトはこれらふたつのタイプのオブジェクトの中間にあると考えてもさして一般性を失わない。したがって、 k は $-1 \leq k \leq 1$ として大抵の場合をカバーでき、式27が一般に減少関数であることが保障される。

【0098】

なお、式34からわかるように、 r は画像の解像度に影響されること、すなわち r は 2 m に比例することに注意すべきである。このために [1. 4. 1] においてファクター 2 m を導入した。

【0099】

[1. 4. 3] η の動的決定

パラメータ η も同様の方法で自動決定できる。はじめに $\eta = 0$ とし、最も細かい解像度における最終的な写像 $f(n)$ 及びエネルギー $C(n)$ f を計算する。つづいて、 η をある値 $\Delta \eta$ だけ増加させ、再び最も細かい解像度における最終写像 $f(n)$ 及びエネルギー $C(n)$ f を計算し直す。この過程を最適値が求まるまで続ける。 η は写像の剛性を示す。次式の重みだからである。

【0100】

【数35】

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式35)

η が0のとき、 $D(n)f$ は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 η が非常に大きな値のとき、 $D(n)f$ は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 η の値が0から次第に増えるとき、後述のごとく $C(n)f$ は徐々に減少する。しかし η の値が最適値を越えると、図4に示すとおり、エネルギーは増加し始める。同図のX軸は η 、Y軸は Cf である。

【0101】

この方法で $C(n)f$ を最小にする最適な η の値を得ることができる。しかし、 λ の場合に比べていろいろな要素が計算に影響する結果、 $C(n)f$ は小さく揺らぎながら変化する。 λ の場合は、入力が微小量変化するたびに副写像を1回計算しなおすだけだが、 η の場合はすべての副写像が計算しなおされるためである。このため、得られた $C(n)f$ の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つかれば、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【0102】

[1.5] スーパーサンプリング

画素間の対応関係を決定する際、自由度を増やすために、 $f(m, s)$ の値域を $R \times R$ に拡張することができる（Rは実数の集合）。この場合、終点画像の画素の輝度が補間され、非整数点、

【0103】

【数36】

$$V(q_{f^{(m,s)}(i,j)}^{(m,s)}) \quad (式36)$$

における輝度を持つ $f(m, s)$ が提供される。つまりスーパーサンプリングが

行われる。実験では、 $f(m, s)$ は整数及び半整数値をとることが許され、

【数37】

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)}) \quad (式37)$$

は、

【数38】

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2 \quad (式38)$$

によって与えられた。

【0104】

【1. 6】各画像の画素の輝度の正規化

始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー $C(m, s)$ f が大きくなりすぎ、正しい評価がしづらいためである。

【0105】

例えば、図20(a)と図20(b)に示すように人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふたつの顔の間の副写像を計算するために、まず副写像を正規化する。すなわち、最も暗い画素の輝度を0、最も明るいそれを255に設定し、他の画素の輝度は線形補間によって求めておく。

【0106】

【1. 7】インプリメンテーション

始点画像のスキャンに従って計算がリニアに進行する帰納的な方法を用いる。始めに、1番上の左端の画素 $(i, j) = (0, 0)$ について $f(m, s)$ の値を決定する。次に i を1ずつ増やしながら各 $f(m, s) (i, j)$ の値を決定する。 i の値が画像の幅に到達したとき、 j の値を1増やし、 i を0に戻す。以降、始点画像のスキャンに伴い $f(m, s) (i, j)$ を決定していく。すべての点について画素の対応が決まれば、ひとつの写像 $f(m, s)$ が決まる。

【0107】

ある $p(i, j)$ について対応点 $qf(i, j)$ が決まれば、つぎに $p(i, j+1)$ の対応点 $qf(i, j+1)$ が決められる。この際、 $qf(i, j+1)$ の位置は全単射条件を満たすために、 $qf(i, j)$ の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに $(0, 0)$ が最も優先される状態がつづくと、求められる最終の写像に余計な偏向が加わる。本前提技術ではこの状態を回避するために、 $f(m, s)$ を以下の方法で決めていく。

【0108】

まず $(s \bmod 4)$ が 0 の場合、 $(0, 0)$ を開始点とし i 及び j を徐々に増やしながら決めていく。 $(s \bmod 4)$ が 1 の場合、最上行の右端点を開始点とし、 i を減少、 j を増加させながら決めていく。 $(s \bmod 4)$ が 2 のとき、最下行の右端点を開始点とし、 i 及び j を減少させながら決めていく。 $(s \bmod 4)$ が 3 の場合、最下行の左端点を開始点とし、 i を増加、 j を減少させながら決めていく。解像度が最も細かい第 n レベルには副写像という概念、すなわちパラメータ s が存在しないため、仮に $s = 0$ 及び $s = 2$ であるとしてふたつの方向を連続的に計算した。

【0109】

実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補 $(k, 1)$ の中からできる限り全単射条件を満たす $f(m, s)(i, j)$ ($m = 0, \dots, n$) の値を選んだ。第 3 の条件を破る候補のエネルギー $D(k, 1)$ には ϕ を掛け、一方、第 1 または第 2 の条件を破る候補には ϕ を掛ける。今回は $\phi = 2$ 、 $\phi = 100000$ を用いた。

【0110】

前述の全単射条件のチェックのために、実際の手続として $(k, 1) = f(m, s)(i, j)$ を決定する際に以下のテストを行った。すなわち $f(m, s)(i, j)$ の相続四辺形に含まれる各格子点 $(k, 1)$ に対し、次式の外積の z 成分が 0 以上になるかどうかを確かめる。

【0111】

【数39】

$$W = \vec{A} \times \vec{B} \quad (式39)$$

ただしここで、

【数40】

$$\vec{A} = \overrightarrow{q_{f^{(m,s)}(i,j-1)}^{(m,s)} q_{f^{(m,s)}(i+1,j-1)}^{(m,s)}} \quad (式40)$$

【数41】

$$\vec{B} = \overrightarrow{q_{f^{(m,s)}(i,j-1)}^{(m,s)} q_{(k,l)}^{(m,s)}}$$

$$(式41)$$

である（ここでベクトルは三次元ベクトルとし、z軸は直交右手座標系において定義される）。もしWが負であれば、その候補についてはD(m, s) (k, l)に ϕ を掛けることによってペナルティを与え、できるかぎり選択しないようにする。

【0112】

図5 (a)、図5 (b) はこの条件を検査する理由を示している。図5 (a) はペナルティのない候補、図5 (b) はペナルティがある候補をそれぞれ表す。隣接画素 (i, j + 1) に対する写像 $f(m, s) (i, j + 1)$ を決定する際、Wのz成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、 $q(m, s) (k, l)$ は隣接する四辺形の境界線を越えるためである。

【0113】

[1. 7. 1] 副写像の順序

インプリメンテーションでは、解像度レベルが偶数のときには $\sigma(0) = 0$ 、
 $\sigma(1) = 1$ 、 $\sigma(2) = 2$ 、 $\sigma(3) = 3$ 、 $\sigma(4) = 0$ を用い、奇数のときは $\sigma(0) = 3$ 、 $\sigma(1) = 2$ 、 $\sigma(2) = 1$ 、 $\sigma(3) = 0$ 、 $\sigma(4) = 3$ を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は4種類であり、s は 0 ~ 3 のいずれかである。しかし、実際には s = 4 に相当する処理を行った。その理由は後述する。

【0114】

[1. 8] 補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面における正方形 $p_{(i, j)} p_{(i+1, j)} p_{(i, j+1)} p_{(i+1, j+1)}$ が終点画像平面上の四辺形 $q_{f(i, j)} q_{f(i+1, j)} q_{f(i, j+1)} q_{f(i+1, j+1)}$ に射影されると仮定する。簡単のため、画像間の距離を 1 とする。始点画像平面からの距離が t ($0 \leq t \leq 1$) である中間画像の画素 $r_{(x, y, t)}$ ($0 \leq x \leq N-1, 0 \leq y \leq M-1$) は以下の要領で求められる。まず画素 $r_{(x, y, t)}$ の位置 (ただし $x, y, t \in \mathbb{R}$) を次式で求める。

【0115】

【数42】

$$\begin{aligned}
 (x, y) = & (1-dx)(1-dy)(1-t)(i, j) + (1-dx)(1-dy)tf(i, j) \\
 & + dx(1-dy)(1-t)(i+1, j) + dx(1-dy)tf(i+1, j) \\
 & + (1-dx)dy(1-t)(i, j+1) + (1-dx)dytf(i, j+1) \\
 & + dxdy(1-t)(i+1, j+1) + dxdytf(i+1, j+1)
 \end{aligned}$$

(式42)

つづいて $r_{(x, y, t)}$ における画素の輝度が次の式を用いて決定される。

【0116】

【数43】

$$\begin{aligned}
 V(r(x, y, t)) = & (1-dx)(1-dy)(1-t)V(p_{(i, j)}) + (1-dx)(1-dy)tfV(q_{f(i, j)}) \\
 & + dx(1-dy)(1-t)V(p_{(i+1, j)}) + dx(1-dy)tfV(q_{f(i+1, j)}) \\
 & + (1-dx)dy(1-t)V(p_{(i, j+1)}) + (1-dx)dytfV(q_{f(i, j+1)}) \\
 & + dxdy(1-t)V(p_{(i+1, j+1)}) + dxdytfV(q_{f(i+1, j+1)})
 \end{aligned}$$

(式43)

ここで d_x 及び d_y はパラメータであり、0 から 1 まで変化する。

【0117】

[1. 9] 拘束条件を課したときの写像

いままでは拘束条件がいっさい存在しない場合の写像の決定を述べた。しかし

、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束条件としたうえで写像を決定することができる。

【0118】

基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、しかる後、写像 f を正確に計算する。

【0119】

まず始めに、始点画像の特定の画素を終点画像の特定の画素に射影し、始点画像の他の画素を適当な位置に射影する大まかな写像を決める。すなわち、特定の画素に近い画素は、その特定の画素が射影される場所の近くに射影されるような写像である。ここで第 m レベルの大まかな写像を $F(m)$ と記述する。

【0120】

大まかな写像 F は以下の要領で決める。まず、いくつかの画素について写像を特定する。始点画像について n_s 個の画素、

【数44】

$$p(i_0, j_0), p(i_1, j_1), \dots, p(i_{n_s-1}, j_{n_s-1})$$

(式44)

を特定するとき、以下の値を決める。

【数45】

$$F^{(n)}(i_0, j_0) = (k_0, l_0), \\ F^{(n)}(i_1, j_0) = (k_1, l_1), \dots,$$

$$F^{(n)}(i_{n_s-1}, j_{n_s-1}) = (k_{n_s-1}, l_{n_s-1})$$

(式45)

始点画像の他の画素の変位量は、 $p(i_h, j_h)$ ($h = 0, \dots, n_s - 1$) の変位に重み付けをして求められる平均である。すなわち画素 $p(i, j)$ は、終点画像の以下の画素に射影される。

【0121】

【数4 6】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{h=n_s-1} (k_h - i_h, l_h - j_h) weight_h(i, j)}{2^{n-m}} \quad (式4 6)$$

ただしここで、

【数4 7】

$$weight_h(i, j) = \frac{1/||(i_h - i, j_h - j)||^2}{total\ weight(i, j)} \quad (式4 7)$$

【数4 8】

$$total\ weight(i, j) = \sum_{h=0}^{h=n_s-1} 1/||(i_h - i, j_h - j)||^2 \quad (式4 8)$$

とする。

【0122】

つづいて、 $F^{(m)}$ に近い候補写像 f がより少ないエネルギーを持つように、その写像 f のエネルギー $D^{(m, s)}(i, j)$ を変更する。正確には、 $D^{(m, s)}(i, j)$ は、

【数4 9】

$$D^{(m, s)}_{(i, j)} = E^{(m, s)}_{0(i, j)} + \eta E^{(m, s)}_{1(i, j)} + \kappa E^{(m, s)}_{2(i, j)} \quad (式4 9)$$

である。ただし、

【数5 0】

$$E^{(m, s)}_{2(i, j)} = \begin{cases} 0, & \text{if } ||F^{(m)}(i, j) - f^{(m, s)}(i, j)||^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ ||F^{(m)}(i, j) - f^{(m, s)}(i, j)||^2, & \text{otherwise} \end{cases} \quad (式5 0)$$

であり、 $\kappa, \rho \geq 0$ とする。最後に、前述の写像の自動計算プロセスにより、 f を完全に決定する。

【0123】

ここで、 $f^{(m, s)}(i, j)$ が $F^{(m)}(i, j)$ に十分近いとき、つまりそれ

らの距離が、

【数51】

$$\left[\frac{\rho^2}{2^{2(n-m)}} \right] \quad (\text{式51})$$

以内であるとき、 $E_2(m, s)(i, j)$ が 0 になることに注意すべきである。そのように定義した理由は、各 $f(m, s)(i, j)$ が $F(m)(i, j)$ に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

【0124】

【2】具体的な処理手順

【1】の各要素技術による処理の流れを説明する。

図6は前提技術の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い (S1)、つづいて始点画像と終点画像のマッチングをとる (S2)。ただし、S2は必須ではなく、S1で得られた画像の特徴をもとに画像認識などの処理を行ってもよい。

【0125】

図7は図6のS1の詳細を示すフローチャートである。ここではS2で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い (S10)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い (S11)、一連の終点階層画像を得る。ただし、S10とS11の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

【0126】

図8は図7のS10の詳細を示すフローチャートである。もとの始点画像のサイズは $2^n \times 2^n$ とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ m を n にセットする (S100)。つづいて第 m レベルの画像 $p(m, 0), p(m, 1), p(m, 2), p(m, 3)$ から特異点フィルタを用いて特異点を検出し (S101)、そ

それぞれ第 $m-1$ レベルの画像 $p(m-1, 0)$ 、 $p(m-1, 1)$ 、 $p(m-1, 2)$ 、 $p(m-1, 3)$ を生成する (S102)。ここでは $m=n$ であるため、 $p(m, 0) = p(m, 1) = p(m, 2) = p(m, 3) = p(n)$ であり、ひとつの始点画像から 4 種類の副画像が生成される。

【0127】

図9は第 m レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の数値は各画素の輝度を示す。同図の $p(m, s)$ は $p(m, 0) \sim p(m, 3)$ の 4 つの画像を象徴するもので、 $p(m-1, 0)$ を生成する場合には、 $p(m, s)$ は $p(m, 0)$ であると考える。[1. 2] で示した規則により、 $p(m-1, 0)$ は例えば同図で輝度を記入したブロックについて、そこに含まれる 4 画素のうち「3」、 $p(m-1, 1)$ は「8」、 $p(m-1, 2)$ は「6」、 $p(m-1, 3)$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2m-1 \times 2m-1$ になる。

【0128】

つづいて m をデクリメントし (図8の S103)、 m が負になっていないことを確認し (S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0 レベルの副画像が生成された時点で S10 が終了する。第0 レベルの副画像のサイズは 1×1 である。

【0129】

図10は S10 によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが 4 つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7の S11 にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6の S1 による処理が完了する。

【0130】

前提技術では、図6の S2 に進むためにマッチング評価の準備をする。図11 はその手順を示している。同図のごとく、まず複数の評価式が設定される (S30)。[1. 3. 2. 1] で導入した画素に関するエネルギー $C(m, s) f$ と

[1. 3. 2. 2] で導入した写像の滑らかさに関するエネルギー $D (m, s)$ f がそれである。つぎに、これらの評価式を統合して総合評価式を立てる (S 3 1)。[1. 3. 2. 3] で導入した総エネルギー $\lambda C (m, s)$ $f + D (m, s)$ f がそれであり、[1. 3. 2. 2] で導入した η を用いれば、

$$\sum \sum (\lambda C (m, s) (i, j) + \eta E_0 (m, s) (i, j) + E_1 (m, s) (i, j)) \quad (\text{式52})$$

となる。ただし、総和は i, j についてそれぞれ $0, 1 \dots, 2m-1$ で計算する。以上でマッチング評価の準備が整う。

【0131】

図12は図6のS2の詳細を示すフローチャートである。[1]で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしでとられる。画像間の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優れたものになる。

【0132】

図12のごとく、まず係数パラメータ η を0、レベルパラメータ m を0に設定する (S20)。つづいて、始点階層画像中の第 m レベルの4つの副画像と終点階層画像中の第 m レベルの4つの副画像のそれぞれの間でマッチングを計算し、それぞれ全单射条件を満たし、かつエネルギーを最小にするような4種類の副写像 $f (m, s) (s = 0, 1, 2, 3)$ を求める (S21)。全单射条件は [1. 3. 3] で述べた相続四辺形を用いて検査される。この際、式17、18が示すように、第 m レベルにおける副写像は第 $m-1$ レベルのそれらに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いま $m=0$ であってそれより粗いレベルはないが、この例外的な処理は図13で後述する。

【0133】

一方、同一レベル内における水平的参照も行われる。[1. 3. 3] の式20

のごとく、 $f(m, 3)$ は $f(m, 2)$ に、 $f(m, 2)$ は $f(m, 1)$ に、 $f(m, 1)$ は $f(m, 0)$ に、それぞれ類似するように決める。その理由は、特異点の種類が違っても、それらがもともと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式20からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが良好とみなされる。

【0134】

なお、最初に決めるべき $f(m, 0)$ については同一のレベルで参照できる副写像がないため、式19に示すごとくひとつ粗いレベルを参照する。ただし、実験では $f(m, 3)$ まで求まった後、これを拘束条件として $f(m, 0)$ を一回更新するという手続をとった。これは式20に $s = 4$ を代入し、 $f(m, 4)$ を新たに $f(m, 0)$ とすることに等しい。 $f(m, 0)$ と $f(m, 3)$ の関連度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では [1. 7. 1] に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの関連度を密接に保つ趣旨である。また、処理の開始点に依存する偏向を回避するために、 s の値にしたがって開始点の位置を変える点は [1. 7] で述べたとおりである。

【0135】

図13は第0レベルにおいて副写像を決定する様子を示す図である。第0レベルでは各副画像がただひとつの画素で構成されるため、4つの副写像 $f(0, s)$ はすべて自動的に単位写像に決まる。図14は第1レベルにおいて副写像を決定する様子を示す図である。第1レベルでは副画像がそれぞれ4画素で構成される。同図ではこれら4画素が実線で示されている。いま、 $p(1, s)$ の点Xの対応点を $q(1, s)$ の中で探すとき、以下の手順を踏む。

【0136】

1. 第1レベルの解像度で点Xの左上点a、右上点b、左下点c、右下点dを求める。
2. 点a～dがひとつ粗いレベル、つまり第0レベルにおいて属する画素を探す。図14の場合、点a～dはそれぞれ画素A～Dに属する。ただし、画素A～

Cは本来存在しない仮想的な画素である。

3. 第0レベルすでに求まっている画素A～Dの対応点A'～D'をq(1, s)の中にプロットする。画素A'～C'は仮想的な画素であり、それぞれ画素A～Cと同じ位置にあるものとする。

4. 画素Aの中の点aの対応点a'が画素A'の中にあるとみなし、点a'をプロットする。このとき、点aが画素Aの中で占める位置（この場合、右下）と、点a'が画素A'の中で占める位置が同じであると仮定する。

5. 4と同様の方法で対応点b'～d'をプロットし、点a'～d'で相続四辺形を作る。

6. 相続四辺形の中でエネルギーが最小になるよう、点xの対応点x'を探す。対応点x'の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図14の場合、4つの画素がすべて候補になる。

【0137】

以上がある点xの対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第2レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図3に示すように画素A'～D'の間隔が空いていく状況が発生する。

【0138】

こうして、ある第mレベルの4つの副写像が決まれば、mをインクリメントし（図12のS22）、mがnを超えていないことを確かめて（S23）、S21に戻る。以下、S21に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後にS21に戻ったときに第nレベルの写像f(n)を決める。この写像は $\eta = 0$ に関して定まったものであるから、f(n) ($\eta = 0$)と書く。

【0139】

つぎに異なる η に関する写像も求めるべく、 η を $\Delta\eta$ だけシフトし、mをゼロクリアする（S24）。新たな η が所定の探索打ち切り値 η_{max} を超えていないことを確認し（S25）、S21に戻り、今回の η に関して写像f(n) ($\eta = \Delta\eta$)を求める。この処理を繰り返し、S21でf(n) ($\eta = i\Delta\eta$) ($i = 0, 1, \dots$)を求めていく。 η が η_{max} を超えたときS26に進み、後述の方

法で最適な $\eta = \eta_{opt}$ を決定し、 $f(n)$ ($\eta = \eta_{opt}$) を最終的に写像 $f(n)$ とする。

【0140】

図15は図12のS21の詳細を示すフローチャートである。このフローチャートにより、ある定まった η について、第mレベルにおける副写像が決まる。副写像を決める際、前提技術では副写像ごとに最適な λ を独立して決める。

【0141】

同図のごとく、まず s と λ をゼロクリアする (S210)。つぎに、そのときの λ について (および暗に η について) エネルギーを最小にする副写像 $f(m, s)$ を求め (S211)、これを $f(m, s)$ ($\lambda = 0$) と書く。異なる λ に関する写像も求めるべく、 λ を $\Delta\lambda$ だけシフトし、新たな λ が所定の探索打切り値 λ_{max} を超えていないことを確認し (S213)、S211に戻り、以降の繰り返し処理で $f(m, s)$ ($\lambda = i\Delta\lambda$) ($i = 0, 1, \dots$) を求める。 λ が λ_{max} を超えたとき S214 に進み、最適な $\lambda = \lambda_{opt}$ を決定し、 $f(m, s)$ ($\lambda = \lambda_{opt}$) を最終的に写像 $f(m, s)$ とする (S214)。

【0142】

つぎに、同一レベルにおける他の副写像を求めるべく、 λ をゼロクリアし、 s をインクリメントする (S215)。 s が 4 を超えていないことを確認し (S216)、S211に戻る。 $s = 4$ になれば上述のごとく $f(m, 3)$ を利用して $f(m, 0)$ を更新し、そのレベルにおける副写像の決定を終了する。

【0143】

図16は、ある m と s について λ を変えながら求められた $f(m, s)$ ($\lambda = i\Delta\lambda$) ($i = 0, 1, \dots$) に対応するエネルギー $C(m, s)$ f の挙動を示す図である。[1.4] で述べたとおり、 λ が増加すると通常 $C(m, s)$ f は減少する。しかし、 λ が最適値を超えると $C(m, s)$ f は増加に転じる。そこで本前提技術では $C(m, s)$ f が極小値をとるときの λ を λ_{opt} と決める。同図のように $\lambda > \lambda_{opt}$ の範囲で再度 $C(m, s)$ f が小さくなっていても、その時点ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 λ_{opt} は副写像ごとに独立して決めていき、最後に $f(n)$ に

についてもひとつ定まる。

【0144】

一方、図17は、 η を変えながら求められた $f(n)$ ($\eta = i \Delta \eta$) ($i = 0, 1, \dots$) に対応するエネルギー $C(n)$ f の挙動を示す図である。ここでも η が増加すると通常 $C(n)$ f は減少するが、 η が最適値を超えると $C(n)$ f は増加に転じる。そこで $C(n)$ f が極小値をとるときの η を η_{opt} と決める。図17は図4の横軸のゼロ付近を拡大した図と考えてよい。 η_{opt} が決まれば $f(n)$ を最終決定することができる。

【0145】

以上、本前提技術によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先駆的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

【0146】

なお、本前提技術について次のような変形技術も考えられる。

(1) 前提技術では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の2枚の画像間のマッチングをとる場合全般に利用できる。

【0147】

たとえば2枚の画像間で、画素の輝度の差に関するエネルギー E_0 と画素の位置的なずれに関するエネルギー E_1 のふたつを評価式とし、これらの線形和 $E_t = \alpha E_0 + E_1$ を総合評価式とする。この総合評価式の極値付近に注目して α を自動決定する。つまり、いろいろな α について E_t が最小になるような写像を求める。それらの写像のうち、 α に関して E_1 が極小値をとるときの α を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

【0148】

これ以外にも評価式の設定にはいろいろな方法があり、例えば $1/E_1$ と $1/E_2$ のように、評価結果が良好なほど大きな値をとるものを探用してもよい。総合評価式も必ずしも線形和である必要はなく、 n 乗和 ($n = 2, 1/2, -1, -2$ など)、多項式、任意の関数などを適宜選択すればよい。

【0149】

パラメータも、 α のみ、前提技術のごとく α と β のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが3以上の場合はひとつずつ変化させて決めていく。

(2) 本前提技術では、総合評価式の値が最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である $C(m, s)$ f が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によつては単に総合評価式の最小値が最小になるようにパラメータを決めても効果的である。その場合、例えば $\alpha E_0 + \beta E_1$ を総合評価式とし、 $\alpha + \beta = 1$ なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてもよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

(3) 前提技術では各解像度レベルで4種類の特異点に関する4種類の副写像を生成した。しかし、当然4種類のうち1、2、3種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する $f(m, 3)$ だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 s に関する計算量が減る効果がある。

(4) 本前提技術では特異点フィルタによってレベルがひとつ進むと画素が $1/4$ になった。例えば 3×3 で1ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は $1/9$ になる。

(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像を変換する。それ以外の方法として、RGBの各成分について副写像を計算してもよい。

[4] 実験の結果

この前提技術を用いて様々な画像を補間することができる。異なるふたつの視点からの画像を補間すると、中間視点からの画像を生成することができる。これはWWWにおいてきわめて有利である。なぜなら、限られた数の画像から任意の視点画像を生成できるからである。ふたりの人の顔の画像を補間すれば、モーフィングを行うことができる。CTまたはMRIのデータのように三次元オブジェクトの断面データ画像を用いれば、補間の結果ボリュームレンダリングのための三次元オブジェクトの正確な形状を再構築することができる。

【0150】

図18 (a)、図18 (b)、図18 (c) は、写像が中間視点画像を生成するため用いられた場合を示している。ここでは右目画像と左目画像が補間された。図18 (a) は左目から見た始点画像、図18 (b) は右目から見た始点画像、図18 (c) は簡単のために [1. 8] において t の値を 0.5 としたときの中間画像をそれぞれ示している。

【0151】

図19 (a)、図19 (b)、図19 (c)、図19 (d) は、写像を用いて人間の顔のモーフィングを行った場合を示している。ここではふたりの顔を補間した。図19 (a) は始点画像、図19 (b) は終点画像、図19 (c) は終点画像に始点画像を重ね合わせた画像、図19 (d) は $t = 0.5$ のときの中間画像をそれぞれ示している。

【0152】

図20 (a)、図20 (b) は、人と猫の顔の補間に写像を用いた場合を示している。図20 (a) は猫の顔、図20 (b) は人の顔と猫の顔のモーフィング画像をそれぞれ示している。人の顔として図19 (a) の画像を用いた。[1. 6] で説明した輝度の正規化はこの例においてのみ用いられている。

【0153】

図21 (a)、図21 (b)、図21 (c) は、数多くのオブジェクトを含む画像に対して今回的方法を適応した例を示している。図21 (a) は始点画像、図21 (b) は終点画像、図21 (c) は $t = 0.5$ のときの中間画像をそれぞ

れ示す。

【0154】

図22(a)、図22(b)、図22(c)、図22(d)は、MRIによって得られた人間の脳の断面画像を補間するために写像を用いた結果を示している。図22(a)は始点画像、図22(b)は終点画像(上部断面)、図22(c)は $t = 0.5$ の場合の中間画像をそれぞれ示している。また図22(d)は四つの断面画像を用いてボリュームレンダリングを行った結果を斜め方向から見た様子を示す。オブジェクトは完全に不透明であり、補間の結果輝度が51($= 255 \times 0.2$)以上となった画素のみが表示されている。再構築されたオブジェクトは中心付近で垂直にカットされ、その内部が示されている。

【0155】

これらの例において、MRIの画像は 256×256 画素、それ以外の画像は全て 512×512 画素である。画素の輝度は0~255のいずれかの値をとる。[1.3.1]で説明した付加条件は、図21(a)~図21(c)の場合を除き、すべての例で用いられている。これら全ての例において $B0thres = 0.003$ 及び $B1thres = 0.5$ が用いられ、これらの値を変更する必要は全くなかった。各副画像の画素の輝度は図20(a)、図20(b)の場合のみ、正規化された。

[画像データ符号化技術]

以上の前提技術を利用した画像データ符号化技術を説明する。

【0156】

図23は、画像データを符号化する処理を概念的に示す。ここでは画像データがキーフレームとそれ以外の中間フレームからなるものとする。キーフレームは最初から決められていてもよいし、符号化に際して定めてもよい。後者の場合、最初の画像データは、単に複数のフレームからなる通常の動画や医療画像などでよい。

【0157】

ここではキーフレーム決定の過程は不問とし、2枚のキーフレームKF(Key Frame) 200、202があるとする。処理はまず、これらのキーフレーム間で

マッチングをとり、仮想的な中間フレームVIF (Virtual Intermediate Frame) 204を生成する。これらの処理は前提技術で詳述されている。ただし、前提技術ではマッチングの対象となる2枚のキーフレームは、始点画像および終点画像と表現されている。「仮想的な中間フレーム」とは、当初の画像データに実際に含まれていた中間フレーム、すなわち現実の中間フレームではなく、キーフレームからマッチング計算をもとに得られたフレームをいう。

【0158】

つづいて、現実の中間フレームAIF (Actual Intermediate Frame) 206を仮想的な中間フレームVIF 204を利用して符号化する。いま現実の中間フレームAIF 206がふたつのキーフレームKF 200、202の間を $t : (1 - t)$ に内分する位置にあったとすれば、仮想的な中間フレームVIF 204も同様に、キーフレームKF 200、202の間を $t : (1 - t)$ に内分する位置にあると想定して補間する。これを対応点を頂点とする四辺形ベースでトライリニア補間、すなわち x 、 y のふたつの方向に補間する技術は、前提技術の【1.8】で述べた。ここでも同様の技術を利用できるが、トライリニア補間以外の技術を採用してもよい。たとえば、四辺形を考えずに対応点どうしを単純に補間することもできる。

【0159】

現実の中間フレームAIF 206の符号化は、それと仮想的な中間フレームVIF 204の差分画像DI 210をハフマン符号化や算術符号化などのエントロピー符号化、DCT (Discrete Cosine Transform) を用いたJPEGなどの符号化、辞書圧縮、ランレンジス符号化その他の手法で処理することにより実現する。画像データの最終的な符号化データ（以下、単に「符号化画像データ」ともよぶ）は、この中間フレームに関する差分画像の符号化データ（以下、単に「中間フレームの符号化データ」とよぶ）とキーフレームデータの組合せとして得られる。

【0160】

この方法では、符号化側と復号側に同じマッチング機能を実装しておくことにより、キーフレームから同じ仮想的な中間フレームを得る。したがって、キーフ

フレームデータに加えて中間フレームの符号化データを取得すれば、復号側でもとの画像データに戻すことができる。差分画像もハフマン符号化その他の符号化により、効果的に圧縮できる。なお、キーフレーム自体もフレーム内圧縮してもよい。中間フレーム、キーフレームとも、ロスレス、ロッシーいずれの圧縮手法を用いてもよく、その指定を可能とする構成であってもよい。

【0161】

図24は以上の符号化処理を実現する画像データ符号化装置10の構成である。各機能ユニットは、たとえばPC(パーソナルコンピュータ)にCD-ROMなどの記録媒体からロードされるプログラムで実現できる。これは後述の復号装置にもいえる。図25は、画像データ符号化装置10による処理のフローチャートである。

【0162】

画像データ入力ユニット12はネットワーク、ストレージなどから符号化すべき画像データを入力する(S1010)。画像データ入力ユニット12は通信機能、ストレージ制御機能、または自身が画像を撮影する光学機器であってもよい。

【0163】

フレーム分別ユニット14は画像データに含まれるフレームをキーフレームと中間フレームへ分類する(S1012)。キーフレーム検出部16は複数のフレームのうち、直前のフレームとの画像の差が比較的大きいものをキーフレームとして検出する。キーフレームどうしの差を大きくしない配慮であり、符号化効率を改善するためである。キーフレーム検出部16の他の例として、一定間隔でフレームを選択してそれらをキーフレームとしてもよい。その場合、処理が単純になる。分別されたキーフレーム38は中間フレーム生成ユニット18とキーフレーム圧縮ユニット30へ送られる。キーフレーム以外のフレーム、すなわち現実の中間フレーム36は中間フレーム符号化ユニット24へ送られる。

【0164】

キーフレーム圧縮ユニット30はキーフレームを圧縮し、圧縮されたキーフレーム42を符号化データ生成ユニット32へ送り出す。中間フレーム生成ユニッ

ト18のマッチング計算部20は前提技術その他の技術によってキーフレーム間のマッチングをとり(S1014)、フレーム補間部22がマッチング結果をもとに仮想的な中間フレームを生成する(S1016)。生成された仮想的な中間フレーム34は中間フレーム符号化ユニット24へ送られる。

【0165】

中間フレーム符号化ユニット24の比較部26は、仮想的な中間フレーム34と現実の中間フレーム36の差分をとり、差分符号化部28がこの差分を符号化することにより、中間フレームの符号化データ40を生成する(S1018)。中間フレームの符号化データ40は符号化データ生成ユニット32へ送られる。符号化データ生成ユニット32は中間フレームの符号化データ40と圧縮されたキーフレーム42を組み合わせて最終的な符号化画像データを生成、出力する(S1020)。

【0166】

図26は、符号化画像データ300の構成例である。符号化画像データ300は、画像データのタイトルやIDなど、画像データを識別するためのインデックスを格納する画像インデックス領域302と、復号の際に利用すべき情報を格納する参照情報領域304と、キーフレームデータ格納領域306と、中間フレームの符号化データ格納領域308を含み、それらの領域が一体化されている。参照情報として、符号化の方式や圧縮率などの各種パラメータがある。ここではキーフレームとしてKF0、10、20…が、中間フレームの符号化データとしてキーフレーム以外のフレームに関するCDI(Coded Difference Image)1～9、11～19が、それぞれ例示されている。以上が符号化側の処理である。

【0167】

一方、図27は画像データ復号装置100の構成を示す。また、図28は画像データ復号装置100による処理のフローチャートである。画像データ復号装置100は画像データ符号化装置10で得られた符号化画像データをもとの画像データへ復号する。

【0168】

符号化画像データ入力ユニット102は、ネットワーク、ストレージその他か

ら符号化画像データを取得する（S1050）。符号化フレーム分別ユニット104は、符号化画像データに含まれる圧縮されたキーフレーム42と、それ以外の付加データ112を分別する（S1052）。付加データ112には中間フレームの符号化データが含まれる。圧縮されたキーフレーム42はキーフレーム復号ユニット106へ送られ、復号される（S1054）。一方、付加データ112は差分復号ユニット114へ送られ、復号された差分画像116が加算器108へ送られる。

【0169】

キーフレーム38は復号データ生成ユニット110と中間フレーム生成ユニット18へ送られる。中間フレーム生成ユニット18では符号化の際と同じマッチング（S1056）を経て仮想的な中間フレームが生成される（S1058）。仮想的な中間フレーム34は加算器108へ送られ、復号された差分画像116と加算される。加算の結果、現実の中間フレーム36が復号され（S1060）、これが復号データ生成ユニット110へ送られる。復号データ生成ユニット110は、現実の中間フレーム36とキーフレーム38を組み合わせて画像データを復号する（S1062）。

【0170】

以上、実施の形態の画像符号化および復号によれば、画素単位のマッチングで仮想的な中間フレームを生成できるため、画質を維持しながら比較的高い圧縮率が実現する。現実に初期実験でも、全フレームを一様にJPEGで圧縮する場合に比べ、当然ではあるが、同等の主観画質で高い圧縮率が得られている。

【0171】

実施の形態のひとつの変形例として、誤差制御の考え方を導入することができる。すなわち、符号化画像データを復号したとき、もとの画像データとの誤差がある範囲に抑える制御である。誤差の評価式として、ふたつの画像間において位置的に対応しあう画素の輝度値の差分の二乗和がある。この誤差をもとに、中間フレームおよびキーフレームの符号化方式や圧縮率を調節したり、キーフレームの選定をやりなおすことができる。たとえば、ある中間フレームに関する誤差が許容値を超えた場合、その中間フレームの近くにキーフレームを新設したり、そ

の中間フレームをはさむふたつのキーフレームの間隔を狭めることができる。

【0172】

別の変形例として、画像データ符号化装置10と画像データ復号装置100は一体化することができる。その場合、中間フレーム生成ユニット18を中心に構成を共用できる。一体化された画像符号化・復号装置は、画像を符号化してストレージへ格納し、必要なときにそれを復号して表示等することができる。

【0173】

さらに別の変形例として、画像データ符号化装置10は、キーフレームおよび装置外で生成された仮想的な中間フレームを入力するところから処理を開始してもよい。その場合、図24の中間フレーム符号化ユニット24と符号化データ生成ユニット32、および場合によりキーフレーム圧縮ユニット30のみで装置を構成することができる。そのほかにも、いずれの構成を装置外に出すかについては当業者に理解されるごとく相当の自由度があり、それらも実施の形態の変形例とする。

【0174】

同様に、画像データ復号装置100も、装置外で復号されたキーフレームと仮想的な中間フレーム、および中間フレームの符号化データを入力するところから処理を開始してもよい。その場合、図27の差分復号ユニット114、加算器108、および復号データ生成ユニット110のみで装置を構成できる。装置の構成の自由度の高さは符号化の場合と同様である。

【0175】

実施の形態では、画素単位のマッチングを中心に説明した。しかし、本発明の画像データ符号化技術はそれに限られることなく、キーフレーム間の処理から仮想的な中間フレームを求める過程を経るか、それを前処理とする技術全般を含む。キーフレーム間でブロックマッチングがなされてもよいし、その他、仮想的な中間フレームを生成する任意の線形、非線形処理を適用してよい。復号側についても同様のことがいえるが、実装上のひとつのポイントは、符号化側と復号側が原則として同じ方法で得られた仮想的な中間フレームを想定する点にある。ただし、それとても絶対ではなく、復号側が符号化の際の規則を認識して対応するか

、符号化側が復号側の処理を想定して符号化に配慮すればよい。

【図面の簡単な説明】

【図1】 図1 (a) とは図1 (b) は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図1 (c) と図1 (d) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 0)$ の画像、図1 (e) と図1 (f) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 1)$ の画像、図1 (g) と図1 (h) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 2)$ の画像、図1 (i) と図1 (j) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 3)$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図2】 図2 (R) はもとの四辺形を示す図、図2 (A) 、図2 (B) 、図2 (C) 、図2 (D) 、図2 (E) はそれぞれ相続四辺形を示す図である。

【図3】 始点画像と終点画像の関係、および第mレベルと第m-1レベルの関係を相続四辺形を用いて示す図である。

【図4】 パラメータ η とエネルギー C_f の関係を示す図である。

【図5】 図5 (a) 、図5 (b) は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図6】 前提技術の全体手順を示すフローチャートである。

【図7】 図6のS1の詳細を示すフローチャートである。

【図8】 図7のS10の詳細を示すフローチャートである。

【図9】 第mレベルの画像の一部と、第m-1レベルの画像の一部の対応関係を示す図である。

【図10】 前提技術で生成された始点階層画像を示す図である。

【図11】 図6のS2に進む前に、マッチング評価の準備の手順を示す図である。

【図12】 図6のS2の詳細を示すフローチャートである。

【図13】 第0レベルにおいて副写像を決定する様子を示す図である。

【図14】 第1レベルにおいて副写像を決定する様子を示す図である。

【図15】 図12のS21の詳細を示すフローチャートである。

【図16】 ある $f(m, s)$ について λ を変えながら求められた $f(m, s)$ ($\lambda = i \Delta \lambda$) に対応するエネルギーC (m, s) fの挙動を示す図である。

【図17】 η を変えながら求められた $f(n)$ ($\eta = i \Delta \eta$) ($i = 0, 1, \dots$) に対応するエネルギーC (n) fの挙動を示す図である。

【図18】 図18 (a)、図18 (b)、図18 (c) はそれぞれ、あるオブジェクトに関する左目画像、右目画像、前提技術で生成された補間画像をディスプレイ上に表示した中間調画像の写真である。

【図19】 図19 (a)、図19 (b)、図19 (c)、図19 (d) はそれぞれ、ある人物の顔、別の人物の顔、それらの重ね合わせ画像、前提技術で生成されたモーフィング画像をディスプレイ上に表示した中間調画像の写真である。

【図20】 図20 (a)、図20 (b) はそれぞれ、猫の顔、人と猫の顔のモーフィング画像をディスプレイ上に表示した中間調画像の写真である。

【図21】 図21 (a)、図21 (b)、図21 (c) はそれぞれ、多数のオブジェクトが含まれる左目画像、右目画像、前提技術で生成された補間画像をディスプレイ上に表示した中間調画像の写真である。

【図22】 図22 (a)、図22 (b)、図22 (c)、図22 (d) はそれぞれ、MRに関する始点画像、終点画像、前提技術で生成された補間画像、補間画像をもとに生成されたボリュームリンダリング画像をディスプレイ上に表示した中間調画像の写真である。

【図23】 実施の形態による画像データ符号化の概念を示す図である。

【図24】 実施の形態に係る画像データ符号化装置の構成図である。

【図25】 実施の形態に係る画像データ符号化装置による処理を示すフローチャートである。

【図26】 符号化画像データの構成図である。

【図27】 実施の形態に係る画像データ復号装置の構成図である。

【図28】 実施の形態に係る画像データ復号装置による処理を示すフローチャートである。

【符号の説明】

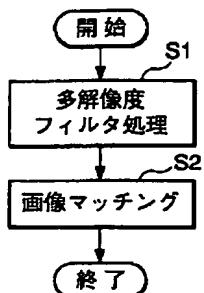
10 画像データ符号化装置
 12 画像データ入力ユニット
 14 フレーム分別ユニット
 16 キーフレーム検出部
 18 中間フレーム生成ユニット
 20 マッチング計算部
 22 フレーム補間部
 24 中間フレーム符号化ユニット
 30 キーフレーム圧縮ユニット
 32 符号化データ生成ユニット
 34 仮想的な中間フレーム
 36 現実の中間フレーム
 38 キーフレーム
 40 中間フレームの符号化データ
 42 圧縮されたキーフレーム
 100 画像データ復号装置
 102 符号化画像データ入力ユニット
 104 符号化フレーム分別ユニット
 106 キーフレーム復号ユニット
 110 復号データ生成ユニット
 112 付加データ
 114 差分復号ユニット
 116 復号された差分画像、すなわち復号された中間フレームの符号化データ
 200、202 キーフレーム
 204 仮想的な中間フレーム
 206 現実の中間フレーム
 210 差分画像
 300 符号化画像データ

302 画像インデックス格納領域

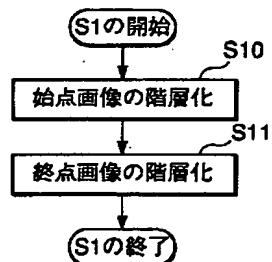
306 符号化中間フレーム格納領域

【書類名】 図面

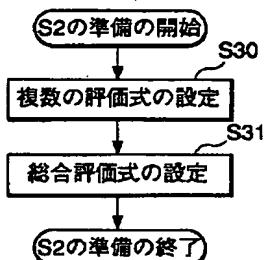
【図1】



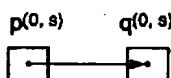
【図2】



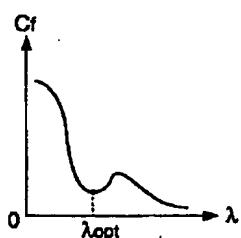
【図3】



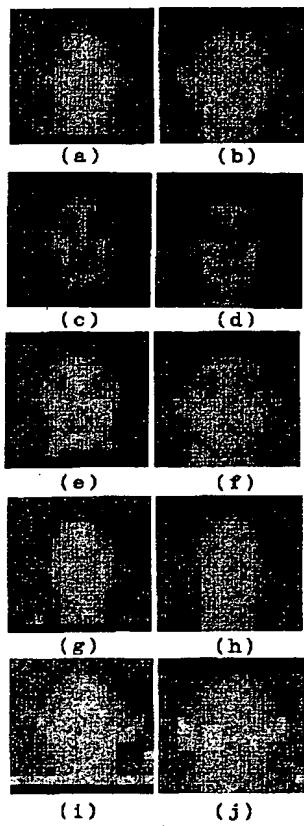
【図4】



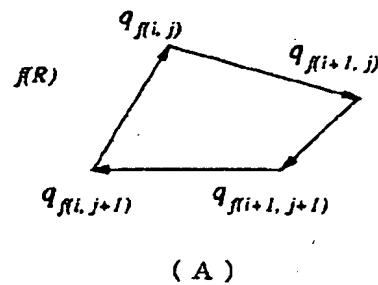
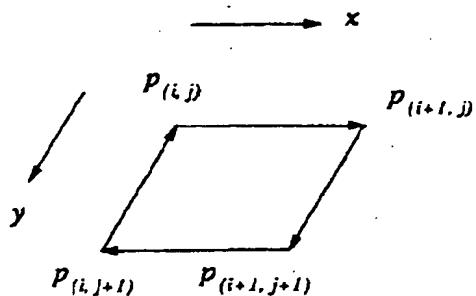
【図5】



【図6】



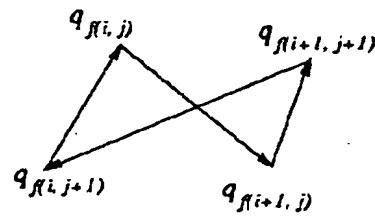
【図7】



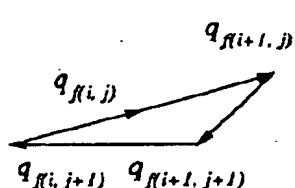
(R)



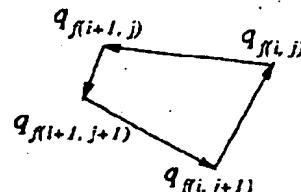
(E)



(B)

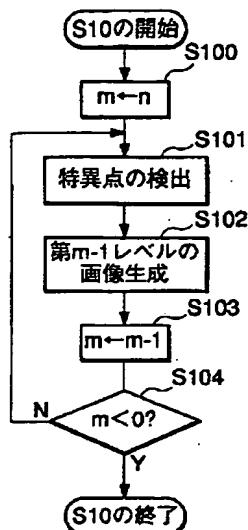


(D)

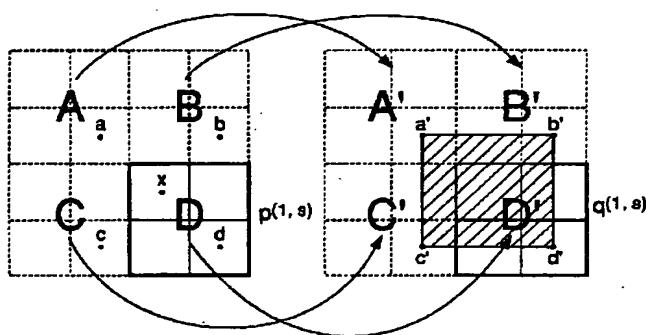


(C)

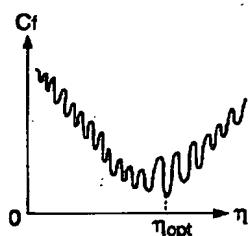
【図8】



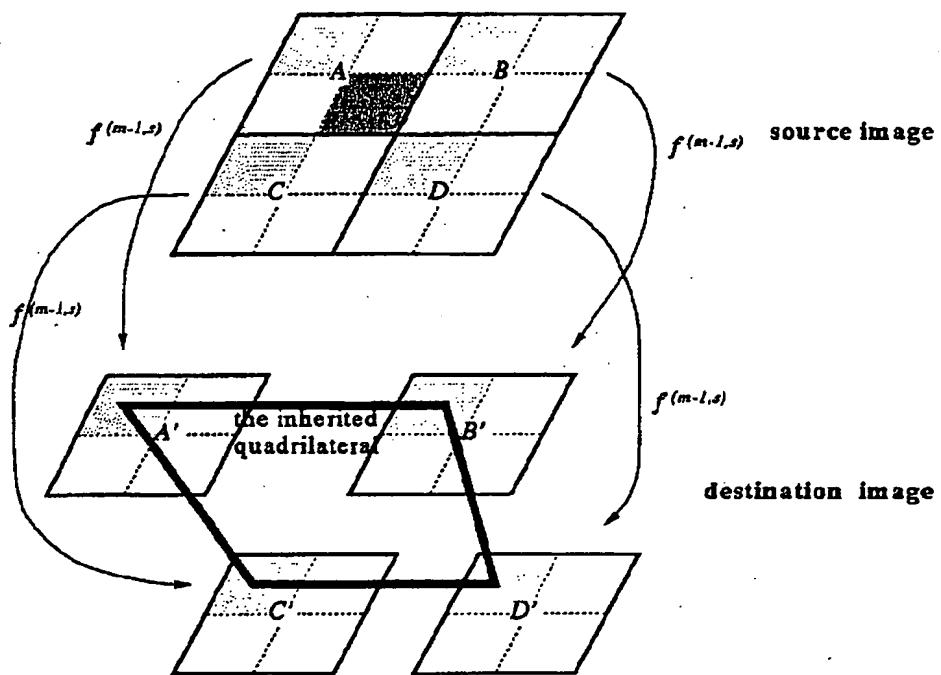
【図9】



【図10】

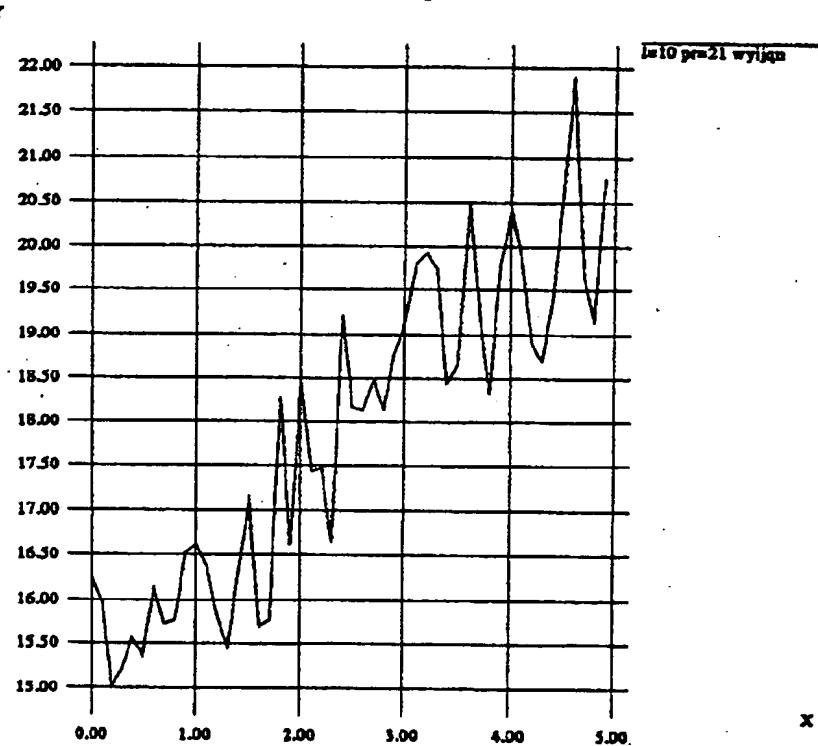


【図11】

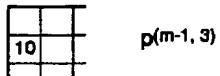
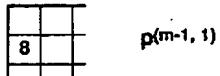
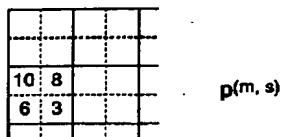


【図12】

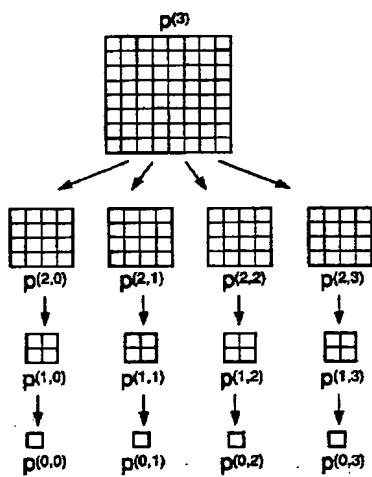
X Graph



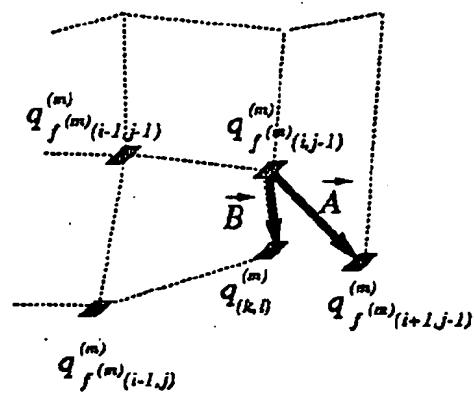
【図13】



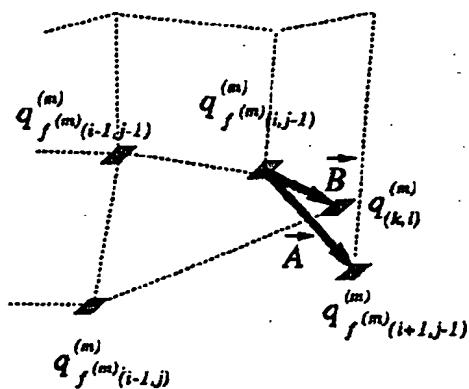
【図14】



【図15】

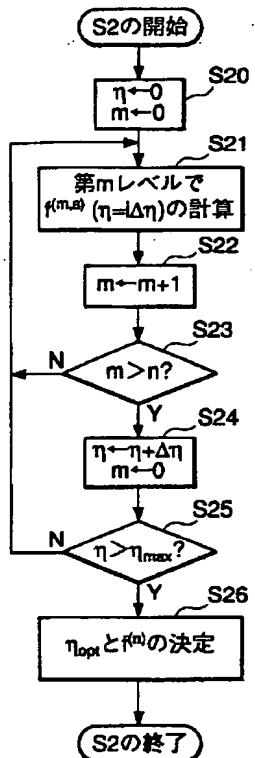


(a)

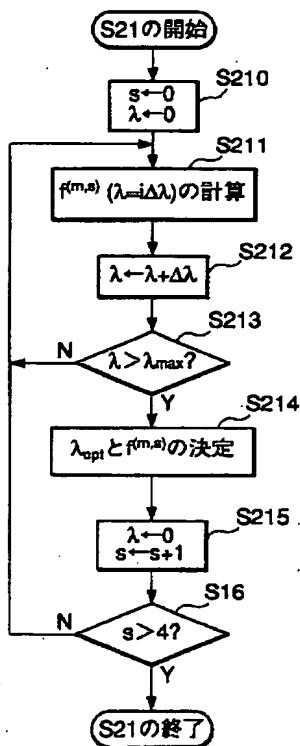


(b)

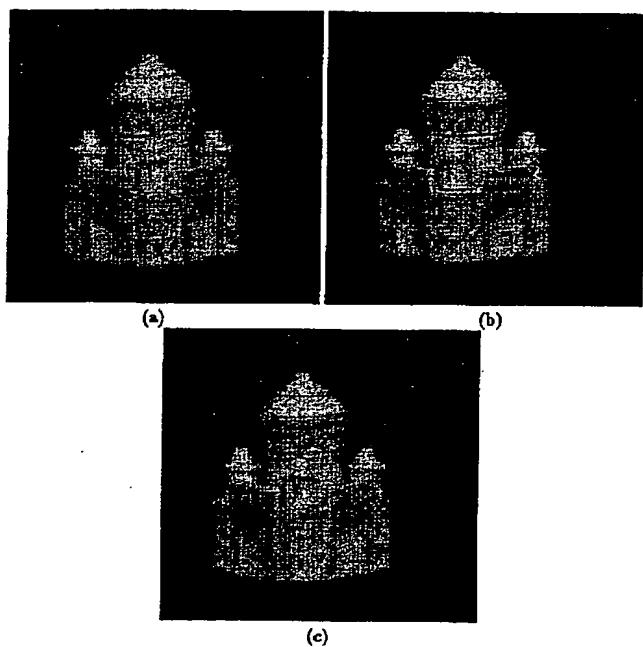
【図16】



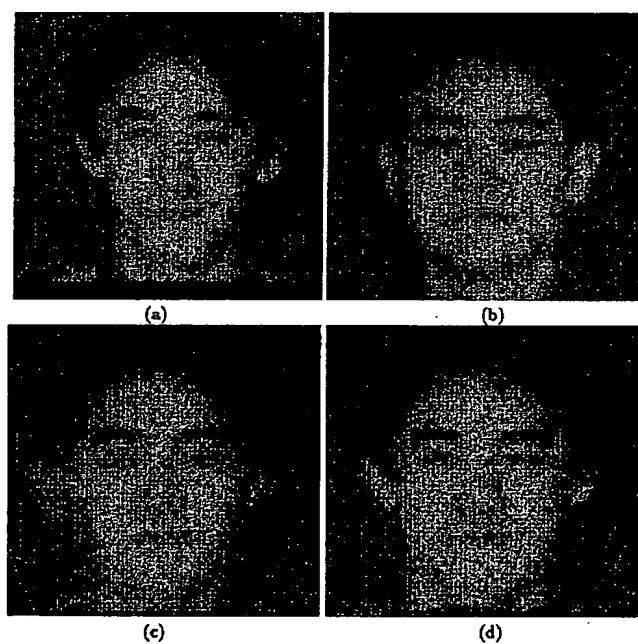
【図17】



【図18】

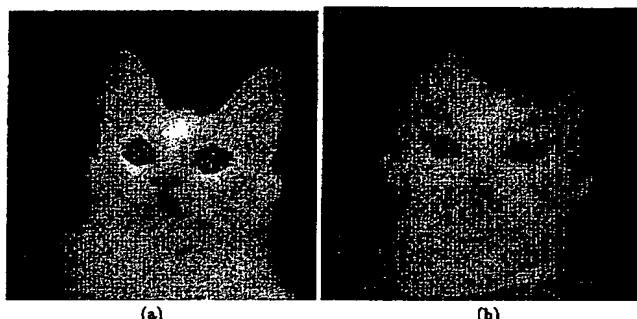


【図19】

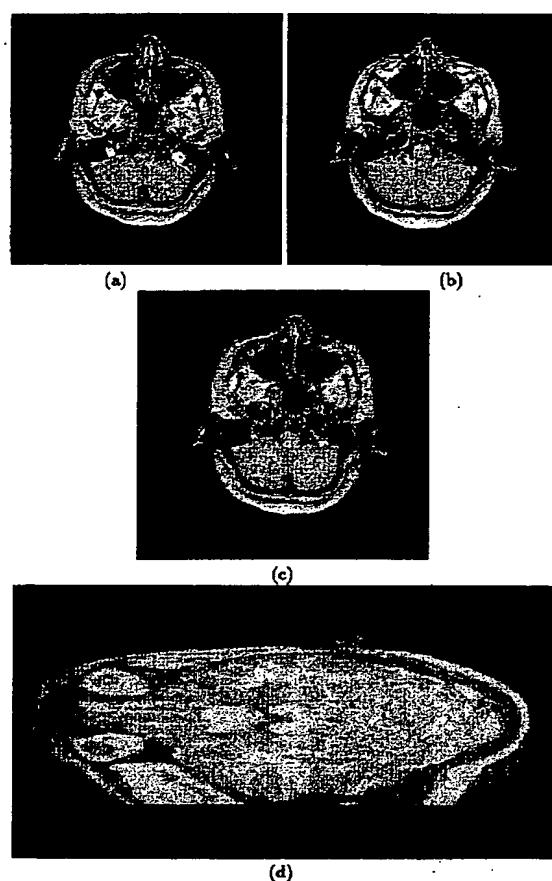


BEST AVAILABLE COPY

【図20】



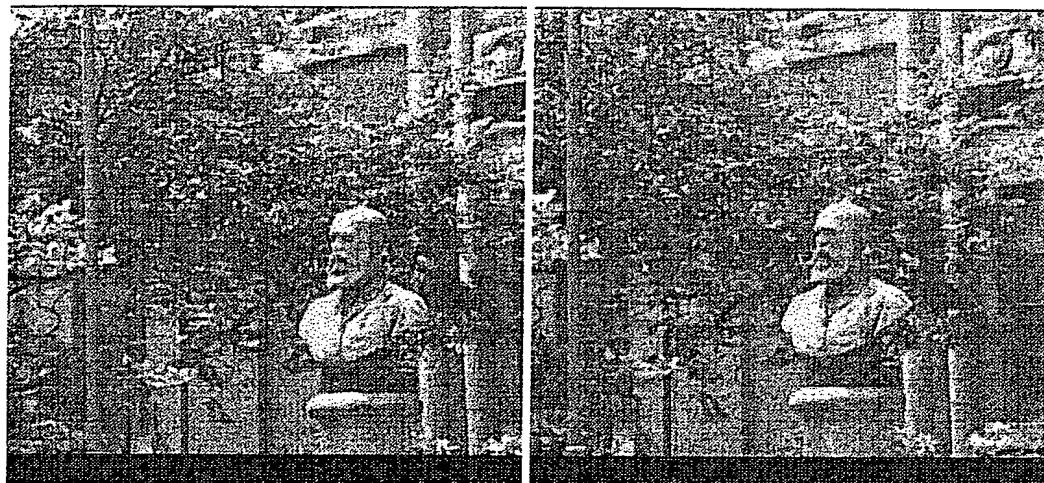
【図21】



BEST AVAILABLE COPY

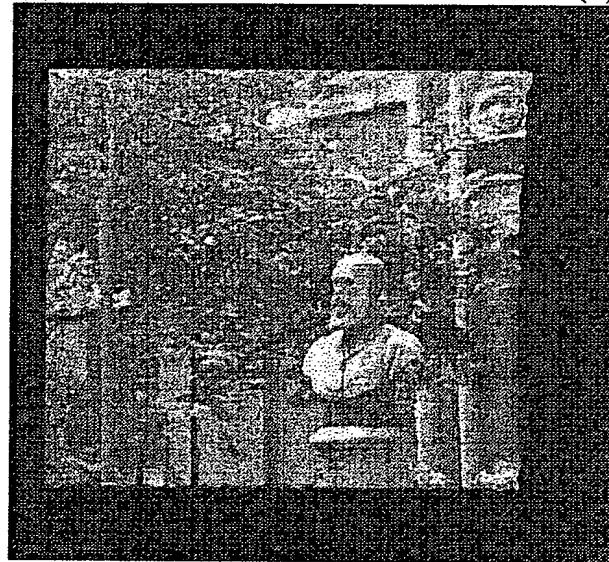
【図22】

カラーステム



(a)

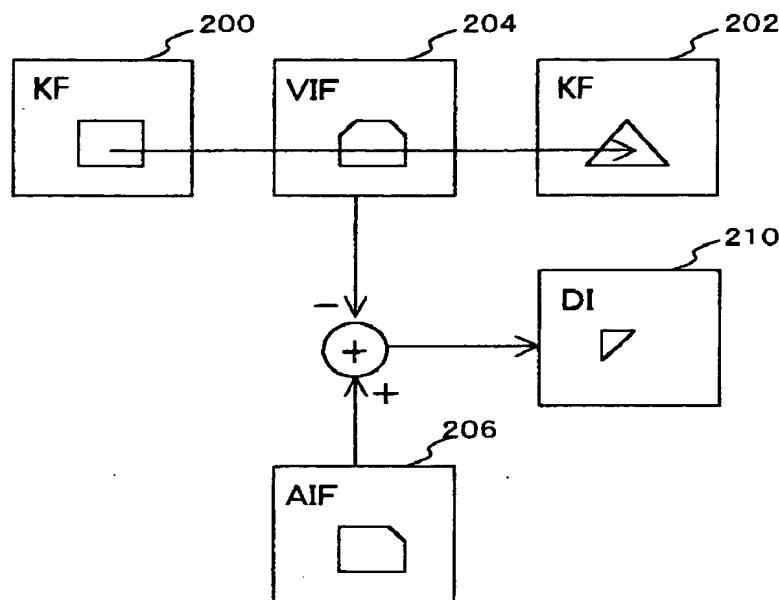
(b)



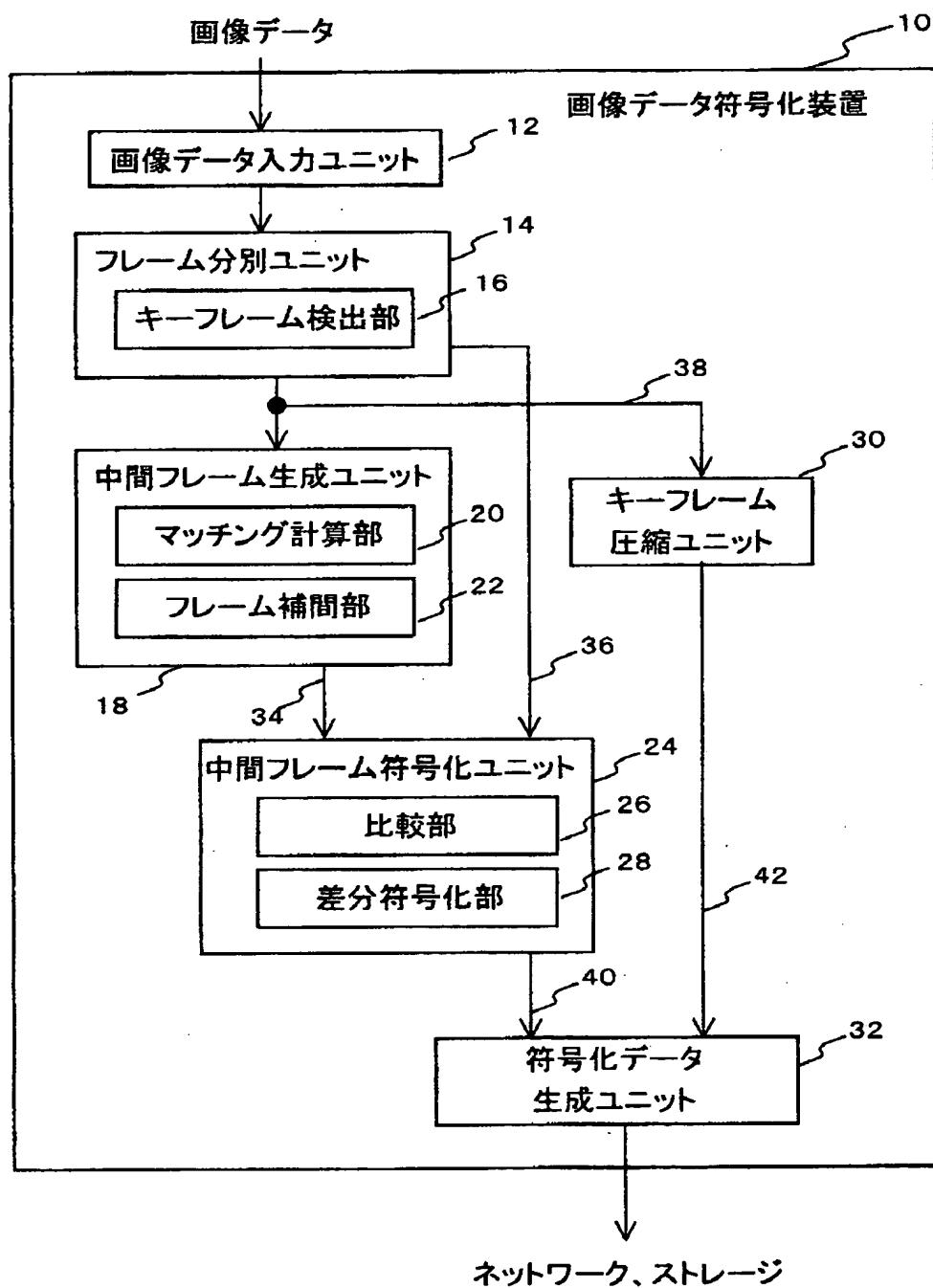
(c)

BEST AVAILABLE COPY

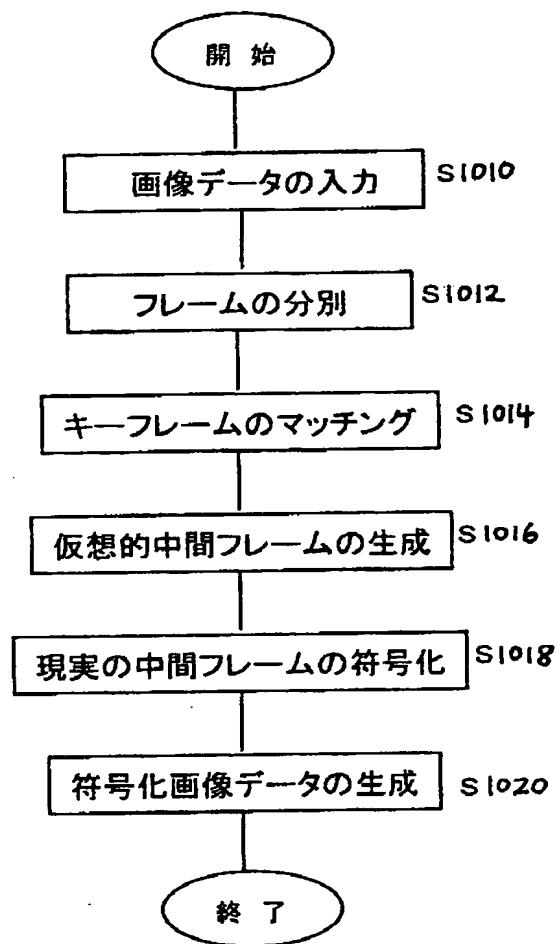
【図23】



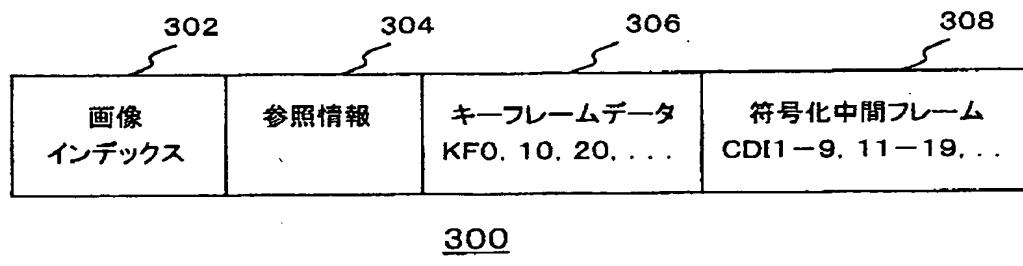
【図24】



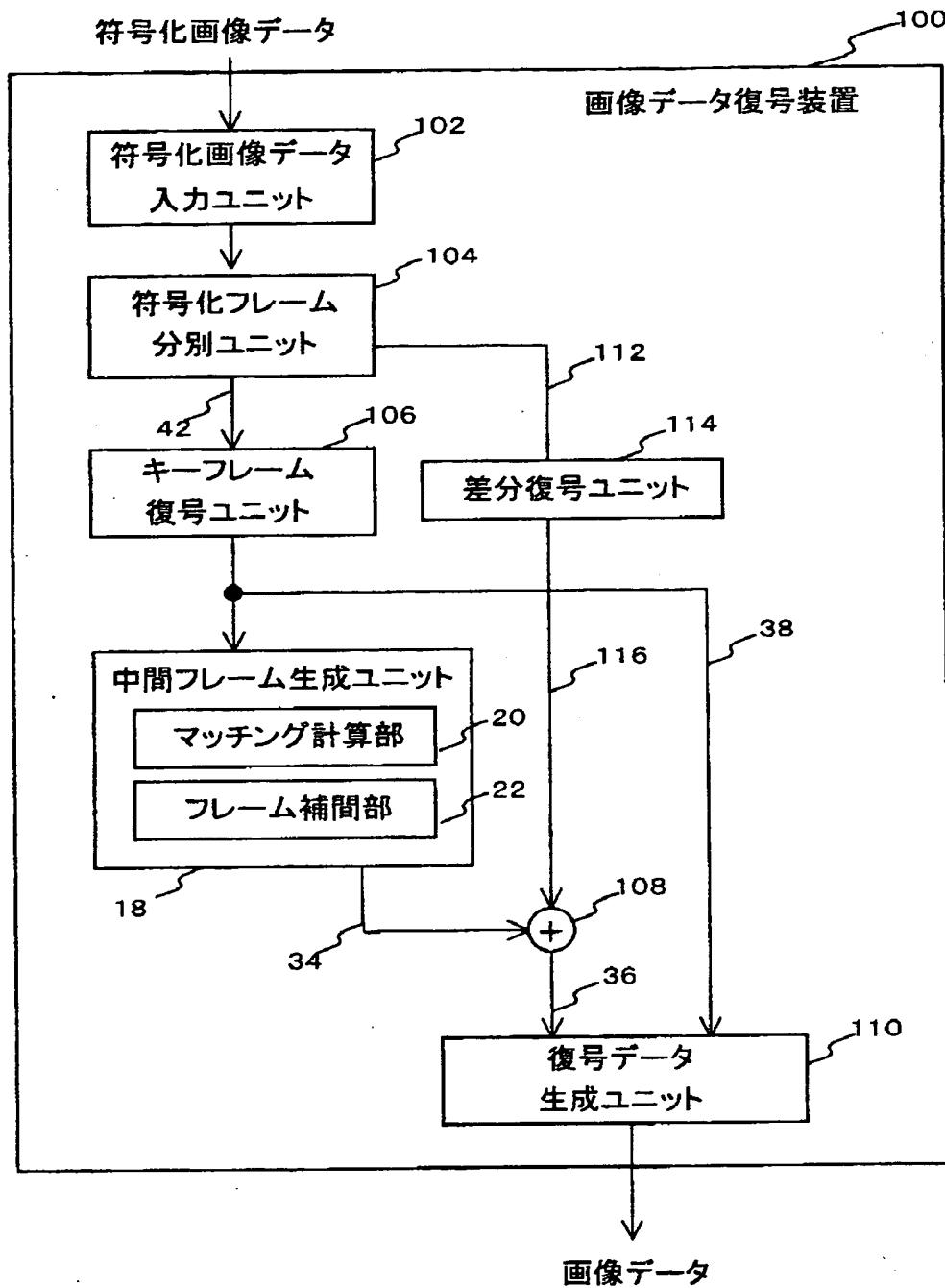
【図25】



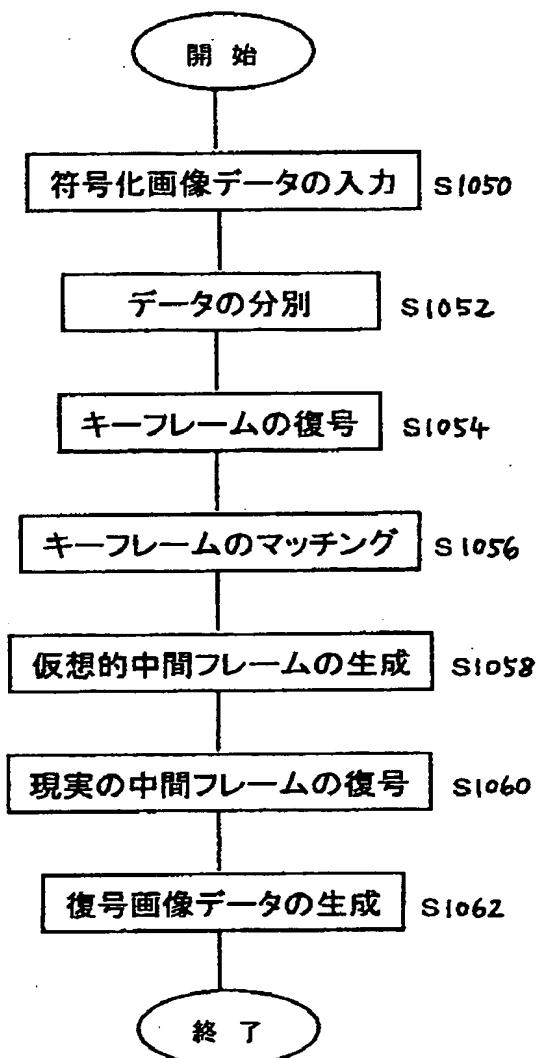
【図26】



【図27】



【図28】



【書類名】 要約書

【要約】

【課題】 画像の精緻化に対し、画像を高効率かつ少ない誤差で符号化する技術が追いつかない状態にある。

【解決手段】 画像データを入力し（S1010）、これをキーフレームとその他の中間フレームへ分別する（S1012）。キーフレーム間でピクセルマッチングを行い（S1014）、マッチング結果を補間することでそれらキーフレーム間の仮想的な中間フレームを生成する（S1016）。この仮想的な中間フレームと実際に当初より存在した現実の中間フレームの差分をとることで、現実の中間フレームを少ない符号量で符号化する（S1020）。

【選択図】 図25

出願人履歴情報

識別番号 [396001980]

1. 変更年月日 1996年 1月26日

[変更理由] 新規登録

住 所 東京都港区麻布十番1丁目7番3号

氏 名 株式会社モノリス